

Can science be made to work?

Most governments are at their wits' end to wring more economic benefit from their investment in research, but only a few seem to be facing up to what needs most urgently to be done.

WHAT does Mr Mikhail Gorbachev have in common with most other political leaders, in both the East and West? The common denominator of politicians' discontent has become the conviction that the economies they run are being deprived of the economic benefits to which a generous regard for the importance of research would normally entitle them. In the pages which follow this week, there are accounts not just of Mr Gorbachev's exhortation to the Party Congress in Moscow, but about the Japanese prime minister's regrets about the quality of university education in Japan and of the French decision, going against the grain, that French companies might usefully participate in the US programme to build a defence against nuclear missiles, not so much for the sake of what is planned but for the technological benefits that might flow from collaboration. In more or less any other week, there would have been a similar chorus of regretful complaint, but the leaders would have been from the United States, from West Germany or, almost certainly, from Britain.

For a common problem, it is natural to look for a common solution. What, in the supposed under-exploitation of science, could this be? The most obvious difficulty is that the common problem has a diversity of causes. The Soviet Union, for example, has a splendid system of secondary education, in stark contrast with that of the United States (to judge from what the US government was saying two years ago, when this cause was briefly fashionable). But Mr Gorbachev is probably right when he now says that the trouble in the Soviet Union is that the machinery of planning, which ostensibly exists to ensure the rational exploitation of resources, is not a help but a hindrance to the creative economic use of research. In the United States, by contrast, where the market is forever sending unmistakable signals to the productive enterprises, it is ever more apparent (in contrast with Japan) that the part of the great American dream that the transition from rags to riches could be accomplished in mere decades by people of sufficient enterprise and quality (not always endearing) has become an excuse for the toleration of amateur bravado in too many parts of the economic machine. Western Europe, not nearly as homogenous as the name suggests, is in a different case again: the trouble is not so much tradition as the prevalent chauvinistic disregard for economic logic, embodied in the belief that even institutions such as the European Community are places in which the participants play a zero-sum game.

Magic

Chauvinism is complicated by impatience. The common theme in what the statesmen are forever saying about the unwon economic benefits of science is that a suitable waving of the right magic wand would put things to rights in a twinkling. In the United States, the hope still lingers that it will be possible to do something about industrial competitiveness in time to affect next year's balance of payments. Mr Gorbachev has the next five years in his sights. What now appears in Western Europe as disappointment with the fruits of science will be bitter anguish ten years from now. Yet, on present form, the message will then be much what it is now. For the truth is that the structural

weaknesses underlying the present weakness of the industrial economies (taking Japan as the yardstick) will not be put right as quickly as people have taken to hoping. Perhaps only China has the timescale right, with its hope of being a modern state by 2025 or thereabouts.

The timescale is important because it determines the choice of policies, among a bewildering variety, that governments should follow. In Britain, there has recently been a great fuss about the sale, to a consortium of US and Italian companies, of a minority stake in Westland helicopters. The argument is that helicopter manufacture is a crucial part of an industrial economy's strategic capability. But nobody would suggest that twenty years from now, Britain will need still to remain self-sufficient in this exotic craft, especially when there are at least three other European companies already busy in the field, and apparently more successful than Westland. (The case of the proposed sale of the Land Rover company, now in public ownership, is different, for that is a successful company in spite of having been badly managed.) The snag, not just in Britain but in the rest of Western Europe, is that the will to make science useful is more often translated into rows over issues such as these than into the building of the institutions from which success might ultimately spring. It is only fair, even to the politicians, to acknowledge that their impatience accurately reflects that of their constituents who, after the title of Mr Kingsley Amis's novel, "want it now".

Solutions

They (that is, we) will have to wait. All of the complaints now being made have their origins in problems that will need decades for solutions. The US high-school system will have to be made good before it can serve what are called societal needs, but it is an unpopular cause to which politicians pay only passing attention. The Soviet Union will either have to become much better at planning or will have to follow China in allowing a simulation of a market economy. In most of Western Europe, the secondary education system is better than that of the United States but less flexible, and thus less likely to throw up young people with a sense of being able to change the world. Britain (in contrast with much of the rest of Europe) will probably be found, some decades from now, to have taken a step backwards by damaging the research enterprise at this crucial time. Japan, curiously enough, may be in the same trouble unless Mr Nakasone's colleagues listen to what he is saying about the temper of academic science.

Europe is best placed quickly to remedy its deficiencies, if only because the framework of the Treaty of Rome allows for the virtually instant creation of a common market. But although Denmark, in last week's referendum, has not run that ship aground by declining to accept the latest community reforms, and despite the moderate success of attempts to stimulate commercial collaboration in technology, the prospects of radical change are as slim as ever. Much the same is true elsewhere. Even Mr Gorbachev's long speech last week may not be remembered for very long. Yet the cultivation of a proper perspective of the hard problems of technology is the only place from which to start. Impatience is the enemy. □



Genetic engineering

Approval for first British virus release experiment

IN advance of the expected publication at the end of this month of the British guidelines for the release of genetically engineered organisms into the environment, approval has been given for the first such release experiments. The experiments, which will involve the release of slightly engineered versions of viruses that are used in the biological control of pests, are said to have been a useful first exercise for the Advisory Committee on Genetic Manipulation (ACGM) that has produced the guidelines and awaits their approval by the Health and Safety Commission.

The guidelines will be just that, and are intended to evolve on a case-by-case basis. It will not be compulsory either to notify ACGM of planned releases or to heed their advice, but the expectation is that nobody will risk bypassing the voluntary procedures. Drawn up by a subcommittee chaired by Sir Robert Williams, but with a membership that has not been made public, the guidelines will put the initial onus on the experimenter to show that adequate safety tests have been, or will be, carried out in advance of release and that there are adequate plans to monitor the release and to abort the experiment if it proves necessary.

Questioned by ecologists at a meeting organized by the Ciba Foundation in London last week, Mr Brian Agar, secretary of ACGM, gave assurances that ecologists were represented on the drafting committee, and Sir Robert emphasized that the proposer of an experiment will be expected to have given consideration to ecological aspects before notifying ACGM.

Half of the meeting was given over to leading British ecologists, Professors John Lawton and Mark Williamson of York University, who offered a perspective from what has been learnt of the results of accidental or deliberate release of animals or plants into novel environments. Drawing particular attention to the inadequacy of ecological modelling to predict with accuracy the outcome of such releases and to a number of disasters that have been documented, Williamson emphasized the need for contingency plans.

Lawton, on the other hand, drew attention to the fact that rigorous screening has so far avoided any serious ecological mishap associated with the several hundred successful applications of biological control of pests. Based on the data for the adaptation of insects to trees newly introduced into Britain, he estimated that the probability of an insect of fastidious diet switching to a newly encountered species

of tree was of the order of 10^{-8} per year. Nevertheless, just as there had been a number of completely unpredictable switches, so the unpredictable has to be reckoned with in the release of genetically engineered organisms. One difficulty is the absence of natural history for such organisms and the difficulty of devising adequate pre-release screening.

What may well be the first release experiment in the United Kingdom will combine a very minor piece of genetic engineering with a major project in biological control. The problem is that of the damage brought about by the moth *Panolis flammea* on the lodgepole pine trees planted in great numbers in the hills of Scotland by the Forestry Commission.

In many areas of Britain the caterpillar stage of the moth is subject to natural control by a baculovirus that infects and kills it. But the virus is not present in the lodgepole pine plantations and so is being introduced: this year several hundred hectares will be sprayed with virus in a full-scale trial of the technique.

But the Natural Environmental Research Council's Institute of Virology in Oxford has plans to improve the effectiveness of the virus by engineering its genes. As a first step in the direction of releasing altered viruses, workers at the institute have engineered a very small change into the virus and are now busy testing its stability and biological activity in the laboratory. The change is carefully designed not to affect of the proteins produced by the virus and is meant merely to serve as a marker by which the virus can be distinguished from its parent.

Dr David Bishop, director of the institute, has already gained approval in principle for release of the virus from ACGM and the Ministry of Agriculture Food and Fisheries as well as support from the Forestry Commission and the Nature Conservancy Council. The first release experiments could take place this summer, says Bishop.

The prediction is that the slightly altered virus will behave just like its natural parent, but the variation will serve as a marker to help test the prediction. If all goes well, the next stage will be to engineer into the virus some self-destruct mechanism for use in conjunction with any of the major modifications that might be tried in the future — such as the introduction of a gene for a caterpillar-killing toxin.

Peter Newmark

Ice-minus bacteria

Further snag and further delay

Washington

A NEW controversy has sprung up around Advanced Genetic Sciences (AGS), the California biotechnology company hoping to begin field tests of a genetically altered bacterium designed to protect crops from frost damage. The field test was to be the first environmental release of genetically altered material, but allegations have surfaced that the company may have released the altered bacteria into the environment during experiments as long as a year ago.

Both the Environmental Protection Agency (EPA), the government regulatory agency that has approved the field test of the bacteria, and the House of Representatives subcommittee that has been overseeing genetic research, have indicated they will investigate the incident.

According to AGS, the experiments in question involved injecting 20 ml of a solution with a low concentration of the altered bacteria into the branches of 45 trees of 6 different species. The trees stand on the asphalt-covered roof of AGS's Oakland headquarters. The roof is surrounded by a 3-foot-high wall, but otherwise is exposed to the environment. The trees were inoculated to test for any pathogenicity from the bacteria.

John Bedbrook, AGS vice-president

and director of research, argues that "since the trees were inoculated in a non-aerosol-producing manner, and the bacteria were contained within the woody plant tissue . . . the conditions of the test were under physical containment" and there was no environmental release.

But EPA believes the roof-top inoculations did constitute environmental release, and that AGS should have notified the agency before proceeding. In fact, AGS did notify EPA about the experiment, but only after the event. The results showed that the bacteria were not pathogenic, and bolstered the AGS application for approval of the field trial.

AGS is hopeful that the current controversy will not further delay the field test of its "ice-minus" version *Pseudomonas syringae* (see *Nature* 319, 254; 1986). The test is for the time being suspended while AGS negotiates with local government officials in Monterey County, California, where the test is to take place, over a suitable site. Objections were raised to the initial site because it is in too populous an area. A further potential complication is a lawsuit brought by the Foundation for Economic Trends, headed by Jeremy Rifkin, that asks for an injunction blocking the trial.

Joseph Palca

Archaeology congress

US plans to boycott boycott

Washington

THE American Council of Learned Societies (ACLS) has added to the headaches of organizers of the World Archaeological Congress by refusing to provide travel funds for any US scientists wishing to attend the meeting. ACLS, a major source of funds for foreign travel, took its decision because of the ban on South African and Namibian scientists from the congress, to be held in Southampton, England, this September.

On 20 February, ACLS wrote to all those requesting travel grants stating that no money would be provided for attending a meeting "where individual scientists have been banned." So far only 29 archaeologists have sought funds from ACLS.

The meeting will take place in spite of the loss of support from the International Union of Prehistoric and Protohistoric Scientists (IUPPS), and despite the resignation of two thirds of the conference's organizing committee (see *Nature* 319, 524; 1986).

The ACLS decision follows a similar move by the Wenner Gren Foundation not to provide travel funds for the meeting. The foundation gave start-up money in 1984 to Peter Ucko, the organizer of the Southampton meeting. Lita Osmundsen, president of the foundation, says a second cheque for \$12,000 had been authorized, but was cancelled when the foundation learned of the ban.

Another possible source for travel money has also been eliminated, at least for the time being. The Society of American Archaeologists (SAA) had intended to apply for a block grant from the National Science Foundation, but changed its mind after learning of the ban.

SAA president Don Fowler says his organization has not issued a policy statement on the Southampton congress, but instead will leave the decision on whether to attend to individual members. SAA's board will meet again in April, when it will reconsider its decision not to seek funds.

US archaeologists are deeply divided over the Southampton congress. At least 45 archaeologists have formed American Archaeologists Against Apartheid and, in a published letter (*Science* 231, 319; 1986), have urged colleagues to attend the meeting as a statement against apartheid.

But others feel that scientific meetings are not the proper place for symbolic political protests. "It threatens every future scientific congress", says Carmel Schrire, an archaeologist from Rutgers who has been encouraging US archaeologists to boycott the Southampton meeting. Schrire is concerned that, while this year's protest may be over apartheid, next year's could be over capital punishment or politi-

tical prisoners in South America, or any divisive international question.

Schrire claims to have commitments from 150 American archaeologists not to attend the Southampton congress, and she maintains that those boycotting the meeting include the most important names in American archaeology.

Both ACLS and the Wenner Gren Foundation will consider funding travel to the alternative congress being held in 1987 in Mainz.

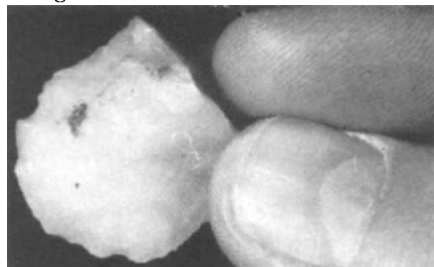
Joseph Palca

Stones turn up

Washington

THE discovery of some 300 stone tools along the Semliki River in the Western Rift Valley in Zaire has anthropologists excitedly speculating on a "fundamental change" in the picture of human development over the past two million years. The tools suggest the presence of early hominids in a part of Africa where they had not previously been found.

The lack of volcanic material in the area where the tools were discovered makes accurate dating difficult. But on the basis of biostratigraphic data and the fossil record, it is estimated that the tools are 2.0–2.5 million years old. Only two sites, in Ethiopia in the Eastern Rift Valley, are thought to be older.



Newly discovered quartzite flake. Dating will provoke controversy.

The tools date from a time when the world climate was becoming drier. The Western Valley was much wetter than the Eastern Valley, perhaps providing a more propitious environment for the survival of early hominids. The sharp edges of the tools suggest that they could have been used in cutting meat and plant stems, as well as woodworking.

The relationship between early hominids and chimpanzees and gorillas will receive closer scrutiny at the Western Rift Valley site, as these animals are not present in the Eastern Rift Valley where early hominids have already been found.

The expedition was led by Noel Boaz of the Virginia Museum of Natural History, John W.K. Harris of the University of Wisconsin–Milwaukee and Alison S. Brooks of George Washington University.

Joseph Palca

Archaeology congress

Mainz meeting fixed for 1987

ARRANGEMENTS for the official congress of the International Union of Pre- and Protohistoric Sciences (IUPPS) have now been completed. At a meeting in Ghent on 21 February, the date of the congress was fixed for the first week in September, 1987. It will be held in Mainz, West Germany, and the organizers are planning for a registration fee of \$20, less than a tenth of that required at the now-unofficial congress to be held this year at Southampton.

The meeting at Ghent was that of a subcommittee of the executive of IUPPS which had been appointed at an earlier meeting in Paris on 17 January to consider replies from the British organizers to the demand for assurances that archaeologists and others from South Africa would be able to attend. In the event, Professor J. Nenquin, the secretary-general of the union, said earlier this week that no communications were received from the British committee, now reconstituted after the resignation of all but two of its members on 8 February.

The choice of the date for the Mainz meeting has been determined by the wish to give the local organizers time to raise the necessary funds, to allow participants from Eastern Europe to arrange travel and, according to Professor Nenquin, to avoid confronting archaeologists with a choice between attendance at Southampton and at Mainz.

Polite recrimination persists among archaeologists, largely on the strength of a document put out on 21 February by the organizers of the Southampton congress. Among other things, this says that the principle of "free interchange between bona fide scholars should prevail under normal circumstances", but that "at present, the circumstances in South Africa and Namibia were not normal".

The statement also says that the meeting in Paris on 17 January was not attended by representatives from developing countries or from Eastern Europe, and that the written opinions of those not present were not admitted. Professor Nenquin says that this procedure was followed while Professor John Evans, then still the chairman of the British National Committee, was in the chair, and that it would have been impossible to take account of all the written opinions by then received.

Nenquin also says that 60 letters have been received from members of the full council of IUPPS, of which 8 oppose the decision to disown the Southampton congress but the majority support their executive committee's decision.

John Maddox

SDI

France changes tack on research

FRANCE is making a swerve, if not an about-turn, in its attitude to the US Strategic Defense Initiative (SDI). A year ago, President Mitterrand said no to French participation in the programme, and proposed Eureka, a project for European cooperation in high-technology projects, as a counterweight. Eureka projects were to be purely civil, but, coincidentally, they were to have covered almost the whole list of its technologies to be tackled in SDI research — until negotiation with European partners modified the concept.

Eureka is still "going well" says its French coordinator, M. Yves Sillard, and could be spending in the first five years as much as \$10,000 million on items such as high-powered lasers, supercomputers and fully automated factories. But going back on its original view, and following its other European partners such as Britain and West Germany, France has now decided to join in SDI, even if M. Mitterrand has not yet graced the initiative with a final Presidential "oui".

Evidence for the change came thick and fast last month, beginning with a speech by defence minister Paul Quilès in which he declared himself "in favour" of French companies signing contracts with the US SDI office in Washington. France may be "far behind the United States in laser research", but it is "the first in Europe", Quilès claimed. France must now "accelerate its research" in lasers to keep up with SDI developments.

The defence minister's thinking was further clarified in a report published a few days after his speech. This was the result of a year's work by a defence ministry special commission on "space weapons", whose president was a French specialist in the interaction of laser beams with gases, plasmas and solid targets, Dr Jean-François Delpeche of the University of Paris Sud at Orsay.

According to the Delpeche commission, the real challenge to France in SDI lies in the Soviet military response — which Delpeche assumes will be to develop its own "guided energy" technology. Therefore, the commission argues, France will need to protect its own independent strategic nuclear weapons from SDI technology. Already, tests are under way at Marcoussis, near Paris, on the protection of French warheads from laser beams by using carbon-fibre nose-cones, the report reveals, but it will be necessary to keep up with SDI laser technology to know how to protect the missiles.

SDI technology could move rapidly, the commission suggests. A full defensive system might not be in place by 2010, but the capacity of US research should not be underestimated. Important discoveries

and innovations could make the impossible "suddenly become realizable".

For a "medium power" like France, the best and cheapest response to SDI will be to develop methods of neutralizing such a system in space (the commission expects it to be "fairly vulnerable"), as well as methods of protecting its own missiles, the report says. France must reinforce its work on continuous and pulsed high-powered lasers to develop studies on laser-target interaction; and must also not forget "third generation nuclear weapons" such as nuclear-pumped X-ray lasers.

In the light of these recommendations Quilès' remarks become clear: France must join SDI to keep up to date with the technology, if only to protect (and de-

velop) its own nuclear force.

The only cautious notes on SDI were left to be struck last week by the French research minister, M. Hubert Curien. He warned that French companies must ensure that technologies developed in partnership with the United States under the SDI banner can be reimported into France, for use in French systems.

It was also important to continue with the Eureka programme, said Curien, because it would be "catastrophic" for Europe if young European researchers gained the impression that all the glamorous research was being done across the Atlantic. For the French research minister, the great danger of SDI for Europe is a brain drain. But many French companies have already established close links with the SDI office in Washington. In SDI, France faces a *fait accompli*. The government has now decided to face it pragmatically.

Robert Walgate

US industrial research

Kodak feels chill winds

Washington

EASTMAN Kodak last week unveiled plans for a research decentralization scheme that will directly link almost 3,000 of its formerly unbridled researchers to business divisions by the end of March. The monolithic research group in Rochester, New York, is being sliced into six units organized according to product lines rather than disciplines, with a seventh 100-man corporate laboratory continuing Kodak's long-term "blue-sky" research. Voices inside and outside the company say the change is long overdue.

The reshuffling will probably be more conceptual than physical. In its previous incarnation, Kodak's research team reported to research administrators, who then conferred with business divisions on product development. Now, individual subsectors are directly responsible to the business divisions. A spokesman said the decentralization should accelerate product development by about a quarter.

Although business-minded managers may demand more focused research projects, most researchers are greeting Kodak's plans with sighs of relief. In fact, the initiative for the re-alignment came from research managers themselves. Obscured communications between the laboratory and the boardroom were putting "a lot of good ideas on the shelf", according to one Kodak engineer. The "cushy research atmosphere," he said, led to "too much wasted effort".

Indeed, wasted effort is one thing Kodak cannot afford. In figures released last month, the \$12,000 million company showed a 64 per cent drop in net earnings, which it blamed on vigorous competition, the strength of the dollar abroad and costs

incurred in patent disputes with Polaroid that have forced Kodak to abandon the promising instant photography market. These grim reports led Kodak to announce "unfortunate but unavoidable" cuts in its worldwide workforce totalling 12,900 people, or 10 per cent. Details of the reductions have not been disclosed, but analysts suspect that research groups may be exempted because of their vital role in Kodak's attempts to foil competitors, the most formidable of which is now Japan's Fuji Film.

Even with the exemption for research, however, an authority at the Rochester Institute of Technology (RIT) noted that Kodak employment is not the pillar of stability it used to be. "They used to be a company that would take an assembly-line worker and turn him into an engineer", said Karen Paul in RIT's business department. But in the past two years, she said, the company has been "scrambling", characterized by some ill-fated acquisitions such as Verbatim, the computer accessories manufacturer.

Kodak harbours hopes that its latest extra-corporate endeavours will not be so unfortunate. Last month the company announced two separate agreements with West Coast biotechnology firms that should shore up its two-year-old life sciences division. Immunex and Kodak have formed a \$30 million joint venture for the development of lymphokine-related products, and Cetus will be farming out its expertise in monoclonal antibodies and DNA probes to expand Kodak's clinical products line. Despite past difficulties, Kodak seems to realize that the road to recovery is paved with good inventions.

Karen Wright

Japanese bioscience

More calls for flexibility

Tokyo

MUCH can be done to strengthen "bioscience" research in Japan, according to the Ministry of Education, Culture and Science (MESC)'s science council. A new series of measures to strengthen bioscience research in the universities was submitted to the minister last week.

The Science Council, consisting of 27 eminent scholars and scientists appointed by the ministry, is the central advisory body on scientific research policy in the universities. The report comes with the science council's official approval, but is actually the work of a 20-member specialist subcommittee set up three years ago. It is expected to exert a considerable influence on MESC's grant policy and strengthen the position of the biosciences — which here, despite an occasional nod towards neurobiology and its possible links with information technology, is taken to mean those sciences using the new techniques of molecular biology.

The report is wide in scope but does not, in its 19 pages, describe how specific proposals should be implemented: its more controversial, and expensive, suggestions are often given as "requiring debate". Overall the council took the modest view that bioscience education and research facilities lag behind those in Europe and the United States. A key point is that bioscience research is by its nature interdisciplinary and rapidly spawns applied technologies which in turn affect basic research. Thus it becomes essential to break down academic barriers — often strong in Japan where departments and the holders of chairs have considerable power and

autonomy. More cooperative research is called for both within the universities and between the universities, attached research institutes and joint-use institutes.

More efforts are needed to ensure adequate supply of research materials; MESC may set up a gene bank as well as computerized databases. More surveys of research trends, both home and abroad, are urged to spot important areas. Plant molecular biology is singled out as a field that has been slow to develop but is likely to be of great importance.

More fellowships for young people and a greater role in their training is urged for research institutes and the joint-use institutes: although the latter often have superior facilities, they have had less say in how graduates are to be trained. A possibility for the longer term is that of setting up a graduate university. More workshops, training and retraining courses are called for, as well as international exchange, both of researchers and of information.

Cooperation with industry is also recognized and a suggestion made that a joint research centre be set up.

Given that the report comes from MESC's key policy body it is likely to affect policy for some time to come and is very good news for molecular biologists. First effects are likely to be seen on grants-in-aid of research. Other effects — on the organization of courses and faculties for example — are likely to take place more slowly. But it is too early to know whether any new money will be made available to support an increased initiative in the biosciences.

Alun Anderson

New research laws miss Nakasone's point

Tokyo

"SOME professors go on repeating the lectures they wrote in their youth until the day they retire", was how Japan's Prime Minister Yoshiro Nakasone uncharitably described university teachers last week. The remarks, prefaced by comments that the level of science in Japan's universities is lower than in the West, were made at a meeting of the House of Representatives budget committee. Nakasone followed up with a controversial suggestion, not unfamiliar to Western researchers, that tenure should be of short duration with evaluation every "five to ten years".

The total security enjoyed by university professors, employed as civil servants by the Ministry of Education, Culture and Science is in no imminent danger of being disturbed, however. Present legislative action to improve the research system is focused on government research institutes. A

new bill, soon to be submitted to the Diet, is designed to open government research institutes to foreign scientists and to allow researchers from private industry to use their facilities.

If the bill becomes law, foreigners will be eligible to occupy any post at government research laboratories — except that of institute chief. Thus an important taboo on allowing foreigners to use public facilities will have been broken.

Changes in patent law are intended to stimulate exchange between industry and government research institutes. At present, rights from joint research, whether in a company laboratory or in a government facility, accrue to the state. Under the new law, rights can be held jointly, or can be transferred entirely to private industry, so that research results will be put to use rather than hoarded by the government.

Alun Anderson

US universities

Academics fear budget cuts

Washington

CLOSER reading of last month's budget seems to have alerted US academics to the seriousness of the consequences for higher education if the federal budget survives the next few months in its present form. Students will be especially badly hit.

The administration's proposals would reduce total spending on direct student support during the coming academic year from \$4,860 million to \$3,840 million. The funds available for Pell grants, which are awarded to students offered a place at a university and able to demonstrate financial need, has already been affected (to the tune of more than \$100 million) by the Gramm-Rudman law; it is now planned to save a further \$359 million by imposing a ceiling of \$2,100 on the value of an award during the coming academic year and by reducing the number of awards, from the 2.7 million now current to 2.1 million next year. A scheme for matching state funds used to make awards to needy students would be eliminated under the proposals.

The administration also proposes to reduce the cost of the student loan programme, but by reorganizing it along lines to be embodied in legislation not yet introduced. At present, the schemes function by the provision of federal subsidies and guarantees to lending institutions. The intention is to save about \$1,000 million a year by requiring that students should pay the full Treasury bill rate during the period of their studies and a higher rate thereafter, while the loan is outstanding.

Although the total amounts involved in institutional support are much smaller, they stand out in the budget proposals by the frequency with which the proposals would eliminate them altogether. Schemes for supporting graduate education, including the National Graduate Fellowships scheme, would be stopped, as would the small schemes for supporting academic libraries and international studies departments at US universities. The scheme for supporting younger institutions with federal grants, the chief source of federal support for university institutions, will remain at roughly the same level but its character will be recast by means of new legislation.

The American Council on Education, which is protesting at the budget resolutions, explains that the uncertainties of the budgetary process in the months ahead imply that it may be harder than in previous years to mobilize support in Congress when individual legislators cannot tell whether amendments to the budget proposals may damage the causes they are seeking to support. □

Publication restraints

Do libel risks prevent publication?

Washington

THE power of threatened libel suits to inhibit the publication of controversial research results was examined last week by the subcommittee on Civil and Constitutional Rights of the US House of Representatives Judiciary Committee. Two researchers at the National Institutes of Health (NIH), Walter Stewart and Ned Feder, described to the subcommittee the difficulties they experienced over a period of more than two years in publishing a study on the frequency of errors in the biomedical literature.

Feder and Stewart were approached by the subcommittee as part of its study of possible changes in the existing libel laws in the United States. One of the committee's chief concerns is that publications with only small circulations, or with small financial resources, such as most scientific journals, are often unwilling to face the high costs of defence against libel suits and so shy away from the publication of controversial material even if they are confident that a suit would fail.

Representative Charles Schumer (Democrat, New York) has proposed legislation (unlikely to be enacted for constitutional reasons) that would allow for factual judgements in libel cases without the award of damages, and which would also ensure that the loser in such a suit would have to pay the legal costs.

Also giving evidence last week was a local newspaper editor from Kentucky who told how his newspaper faces bankruptcy by the defence costs arising from a series of apparently frivolous libel suits, while a former congressional staff member, Andrew Maguire, who alleges he has been libelled in a book by William Morrow called *Poisoning for Profit*, said that he had been unable to obtain redress for lack of financial resources.

Feder and Stewart explained last week that they had embarked on a careful scrutiny of the research papers published over a period of years by a researcher who was later discovered to have forged much of the data on which the papers were based. Feder and Stewart said that they had found the frequency of errors and discrepancies in this sample of research papers to be unexpectedly high. Acknowledging that some of the errors were minor, indicative of haste or carelessness, they said that others were so serious that the fraudster's co-authors should have known them to be false statements.

According to Feder and Stewart, the manuscript was first sent to *Nature* for publication in September 1983, and then sent, with the author's permission, to several of the co-authors of the acknowledged forger. According to Feder and

Stewart, two of the co-authors threatened legal action against *Nature* and the authors. There followed a long negotiation between *Nature* and the authors over changes and revisions of the manuscript; Feder and Stewart told the subcommittee that "at the time, the changes seemed drastic and unnecessary; we were not as familiar then as now with the power of the threat to sue". On 1 February 1985, the authors withdrew their manuscript from consideration by *Nature*.

Feder and Stewart said that the manuscript was next sent to the biology research journal *Cell*. In spite of an initially favourable response, they said, editorial changes were demanded and the editor made it a condition of publication that the authors should indemnify *Cell* against legal costs and liabilities in the event of a lawsuit. When Feder and Stewart agreed to this condition, they said, they were also then asked not to talk to journalists or others about the process of editorial review at *Cell* as a condition of publication. By their account, Feder and Stewart refused, and their study was rejected in November 1985.

According to Feder and Stewart's testimony last week, the manuscript was then sent to a number of other journals for informal comment. Although, they said, more than half the editors concerned responded favourably to the study, publication seemed unlikely, with the result that no formal submission was made. Feder and Stewart said they had had no reply from NIH to a request that the study should be published.

The authors went on to tell the subcommittee last week that most of those individuals to whom the study had been sent had urged that it should be published. Members of the subcommittee seemed to share the view that this "important" work should see the light of day. A revised version of the study is again being considered by *Nature* for publication. **Tim Beardsley**

It is correct that the manuscript referred to was considered for publication in Nature over a period of several months, and that a much revised version has been resubmitted. The original version of the manuscript drew attention to the high frequency of immaterial errors in various research articles, and went on to make inferences about the professional conduct of the forger's colleagues. Strictly speaking, it is not correct that some of the un-named but identifiable co-authors threatened Nature with a libel suit, but they, and their attorneys, argued that the manuscript in its original form and in many later versions would have been unwarrantably damaging. This is true, and has been accepted to be so.

EDITOR, *Nature*

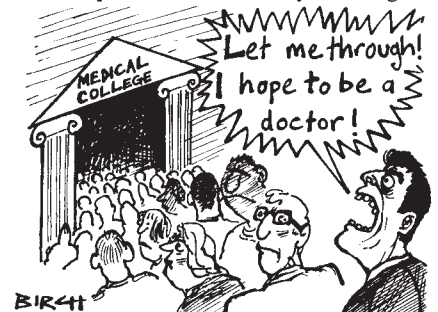
West German physicians

Students still look to medicine

Hamburg

A NEW procedure for the selection of medical students is to be introduced in West Germany with effect from next October. The first stage of the procedure, a multi-centre examination held on 19 February, was a triumph of organization. More than 60,000 students gathered in 800 examinations halls in 226 towns and cities. The 5-hour multiple-choice test was administered on a strictly uniform timetable by 3,000 teachers. Each applicant for a university place to study medicine, dentistry or veterinary medicine is required to undergo the test, and is allowed to attempt it only once. Thereafter, the certificate awarded to each candidate is decisive for his or her future career.

Under the old system, 60 per cent of the medical places were filled by drawing lots



among the applicants. From now on, 45 per cent of medical students will be selected by a weighted combination of results in the new test (45 per cent) and the school-leaving examination, the *Arbitur* (55 per cent), while a further 10 per cent of the places will be reserved for those with the best test results. There will be a 10 per cent quota for foreigners and cases of special hardship such as disability, while 20 per cent of the places will be filled from the 5-year or 6-year waiting list of those who have completed their *Arbitur* but have not so far been able to enter a medical department.

The most controversial feature of the new scheme, giving scope for nepotism say critics, is that the remaining 15 per cent of medical places will be filled at the discretion of the professors.

The irony is that there is as yet no sign that the new system, which has been devised by the *Kulturministerkonferenz* (the regular meeting of the state education ministers) will reduce the numbers of would-be medical students. Even the rapid growth of unemployment among physicians, and the prospect that physicians' unemployment in the 1990s will be as serious as that among school-teachers, seems not yet have filtered through to the schools.

Jürgen Neffe

Soviet science

Gorbachev demands more effort

SOVIET science can expect in the next five years the kind of shake-up which Mr Mikhail Gorbachev's administration has already launched in other sectors of the economy. In his marathon six-and-a-half hour keynote speech to the Twenty-Seventh Party Congress last week, Mr Gorbachev promised Soviet scientists an improved material-technical research base and the creation of "conditions for fruitful activity". In return, he said, the country has the right to expect "discoveries and inventions which ensure genuinely revolutionary changes in the development of technical science and technology". But in many cases, the application of research is being blocked by "the ambitions of individual groups of scientists, departmental hostility towards the inventions of others, and a lack of interest from the production sector".

Particular targets of Mr Gorbachev's criticisms were the research institutes that belong neither to the academy network nor to the higher education sector, but come under various ministries. Last June, at the plenary meeting of the Party Central Committee, Mr Gorbachev pointed out that these institutes had proliferated over the past 20 years more as a means of enhancing the prestige of individual ministries than from considerations of useful and necessary research.

Last year's meeting urged the absorption of such institutes into inter-industrial scientific complexes, cutting across existing ministerial boundaries, but not everyone is in favour of such a move. "We must find out who is opposing it", Mr Gorbachev warned.

The universities and institutes of higher education could also do better, Mr Gorbachev said. More than 35 per cent of Soviet scientists and pedagogic workers are concentrated in the higher education sector, as are almost half of all persons with a Doctor of Sciences degree. Yet only 10 per cent of Soviet research is conducted in the universities. Proposals should be worked out for strengthening ties between the universities and production, in such a way as to take into account the training of the next generation of scientists.

The Academy of Sciences escaped fairly lightly — indeed, at one point, it was even praised for having established a new department of information science and computer technology. But it was instructed fairly sharply to "pay more attention to fundamental research" and to its implementation. That, Mr Gorbachev said, is the "sacred duty" of every scientist, engineer, designer and enterprise director.

Parallel with an intensified research programme, Soviet industry is to be massively re-equipped on a "science-

intensive" basis. Capital investment in industry over the next 10 years will be 200,000 million roubles more than in the preceding decade. Widespread "electronization" (apparently a neologism of Mr Gorbachev's own coining) and integrated automation of industry is planned, also the production on an industrial scale of software for computers and control systems. Furthermore a major drive is envisaged against delays in the construction and modernization programmes which have tied up enormous investment resources and slowed down the rate of technical progress.

Academician Anatolii Aleksandrov, president of the Soviet Academy of Sciences, replied, as it were, on behalf of the scientists. After a virulent polemic against the United States he offered a cautious defence of the research sector. The academy, he said, is concentrating, as Mr Gorbachev would wish, on both fundamental

and applied projects which could be implemented in production, but which, in fact, are not being implemented because of bureaucratic problems.

Soviet science, said Aleksandrov, has now reached a strong position in all the main branches of science — biology (where it has made up its former grave deficits), biochemistry, electronics and micro-processor technology. But because of the problems of reaching an agreement with "far too many ministries", it is difficult to work properly.

"Look", he urged the Congress, "at Comrade V. A. Kotelnikov, sitting here in this hall", who worked out an optical fibre communications system, only to have it rejected by industry because too many ministries would be involved. But now a production complex is being organized, and soon things will be moving in the right direction. For there is no greater grief to a scientist, Aleksandrov concluded, than when he sees the work he has completed, which he knows to be necessary and useful, simply lying inactive on the shelf.

Vera Rich

Polish education

Winning hearts without books

POLAND should establish a network of inter-university laboratories with "unique" equipment, according to Politburo member Tadeusz Porebski. Addressing a plenary meeting of the Central Committee of the Polish United Workers' Party last month, Porebski said that, however inconvenient the use of such facilities might be, there was no other way for Poland to provide the higher education sector with research equipment. The lack of hard currency makes it impossible, he said, to supply all the universities with the equipment they claim to need.

The plenum was convened to formulate party policy on education and youth — what party leader General Wojciech Jaruzelski called "the battle for the minds and hearts of young people". The major part of the discussion was therefore devoted to ideological matters such as the need to strengthen the compulsory university courses in Marxism-Leninism, and the low rate of party membership among students and junior academics. It proved impossible, however, to discuss the ideology of education without dealing with the overall problems facing Polish education.

Higher education, it was admitted, receives insufficient help both from the government departments concerned with the economy and from the production sector which looks to it for a supply of skilled graduates. Investment in education has been increasing over the past five years, reaching, it was claimed, 4.5 per cent of the national income in 1985 (for the entire

educational sector). But an annual expenditure of at least 5 per cent of national income is needed to prevent deterioration of the existing situation, and if the educational system is to be modernized, 7 per cent of national income will be required. The production of textbooks and educational aids (including computers) is way behind schedule, and an "incorrect personnel policy" has led to a lack of job mobility among younger academics, in sharp contrast to the national economy as a whole, where the high job turnover is a source of anxiety to the planners.

Some of the proposals at the congress had both economic and political implications. Postgraduate research in Poland was very considerably curtailed in 1982, and it is now suggested that doctoral studies should be built up by sending more Polish postgraduates to study in other socialist countries, especially the Soviet Union, Czechoslovakia and East Germany. Research funding should give priority to the subjects specified by the integrated Comecon research plan worked out in Moscow last December. And, in spite of repeated pledges by the Minister of Science and Higher Education, Dr Benon Miskiewicz, that no political purge of academics is being planned, the party view, as expressed by General Jaruzelski, was somewhat different. When it comes to the promotion of university lecturers, he said, the essential criteria must include the "ethical and political stance" as well as their scholarly and pedagogic talents.

Vera Rich

Cuban biotechnology

Progress despite isolation

Havana

CUBA's energetic attempts to develop its own biotechnology research base and apply it to human and animal health-care were proudly paraded before 900 delegates from 41 countries last month at the Second Cuban Seminar on Interferon. Some of Cuba's technical achievements — such as the cloning and expression of α_2 -interferon in yeast at high yield — are undeniably impressive, given the total economic blockade by the United States. Visiting US scientists found their Cuban counterparts dedicated and well trained, but some came away with reservations (shared by some Cubans) that a new emphasis on central research planning and pressure to find practical applications might still leave the country short of basic research expertise in a few years' time.

Cuba's decision five years ago to devote a new research centre almost exclusively to the production and purification of interferon — the Centre for Biological Research (CIB) — still seems surprising, given the economic problems the country faces and the relative lack of research in, say, physics. The explanation offered by vice president and minister of education and science José Ramón Fernández, who occupies his spare time compiling books of reprinted articles on biotechnology, is that he and Fidel Castro decided that biosciences could be Cuba's economic salvation next century. He cites Cuba's restricted agriculture (sugar cane is overwhelmingly the largest crop) and poor mineral resources. Fernández says the science budget, as far as it can be distinguished from education, has almost doubled in 5 years to 140 million pesos (1 peso = US\$1.18). The claim that great importance is attached to health care is backed up by impressive statistics: infant mortality and life expectancy are close to the US figures (respectively, 14.3 per thousand, and 74 and 75 for men and women).

One explanation heard for the interferon effort is that it was selected as a "model" to develop the infrastructure for cloning and protein harvesting; other products will follow. Interleukin-2 has been cloned within the past year. But although outside Cuba most now recognize that interferon will probably not be the wonder drug the popular press was proclaiming a few years ago, Fernández seems undaunted.

To emphasize further its seriousness about biotechnology, Cuba is spending 50 million pesos on a new Centre for Genetic Engineering and Biotechnology (now under construction) that will house 150 scientists by this summer and include sterile animal rooms and a P4 containment laboratory. Cuba was a candidate location for one of the two biotechnology research

laboratories to be built by the United Nations Industrial Development Organization (UNIDO), but withdrew its offer to avoid competing with India; it then decided to build one on its own, which is likely to be in operation before the two UNIDO laboratories, now being built in New Delhi and Trieste.

Although it concentrated initially on interferon, CIB has diversified. Under the control of a government agency known as the "Biological Front", CIB is now developing a vaccine for the cattle disease caused by *Clostridium haemolyticum* (antigens have been expressed in *Escherichia coli*), and it plans a new animal research centre.

Many of CIB's scientific staff were trained at the older and much larger National Centre for Scientific Research (CENIC), founded in 1965 on Castro's instructions. More than 8,000 scientists have been trained there before moving on to other institutions; the scientific staff, numbering 1,000, is down from the peak of 1,400 a few years ago. About 50 papers are published each year in international journals, with more in in-house publications. Research at CENIC focuses on public health and the synthesis of drugs and biologicals, as well as some environmental science and corrosion research; the new biotechnology is used wherever possible. But unlike the early days (which were "anarchy", according to director Manuel Limonta) research has recently been placed under much tighter control by the National Academy of Sciences, a division of central government. The process by which research areas are selected is characterized vaguely as a mixture of priority areas handed down from the academy and suggestions funnelled up from the laboratory bench. The deputy prime minister, Fernández, described CENIC as a haven for the basic researcher, but it is clear that a researcher has to obtain the support of a superior before a suggestion will go anywhere. Independent peer review seems to be considered a luxury that Cuba can do without, and researchers and laboratory chiefs make no bones about their preference for research with foreseeable applications.

Many of the health-related research projects are using new techniques of molecular biology. Cuba is perversely proud of the fact that atherosclerosis, normally thought of as a disease of developed countries, is now the major cause of death among males, and efforts are under way to isolate a monoclonal antibody against the apolipoprotein-B receptor to study genetic predisposition to the disease.

Plant biotechnology has recently been recognized as a national priority and is

now a major CENIC activity, much of it focusing on sugar cane. Somaclonal variation is being exploited to produce new varieties which have been introduced on farms and are now found in seven of Cuba's fourteen provinces. Five different varieties have been identified that are resistant to eye spot disease, *Helminthosporium sacchari*, and a protein has been identified in the plant whose quaternary structure seems to correlate with resistance. Other major research efforts concentrate on improving the amino acid composition of yeast for animal feed, extraction and characterization of active substances in plants and ways of using waste from sugar refining, such as producing biogas. Dr Jorge Benitez of the genetics department has constructed stable hybrids of *Saccharomyces cerevisiae* and *Candida utilis* that can be fused (as protoplasts) with the parent strains to give segregants that grow on D-xylose but also show β -glucosidase activity (the enzyme responsible for cellulose breakdown).

None of Cuba's science would be possible without a good educational system, and 220,000 students are in some form of higher education (out of a population of about 10.6 million). Especially talented high-school students are picked out in their teens for fast-track treatment, and Fernández wants to see more fast-track education. Students' exchanges at both undergraduate and postgraduate levels are considered vital, and many researchers were trained in Moscow and other Eastern bloc countries (including Angola) as well as Mexico, Britain and France; about 10,000 Cuban students are now studying abroad. And despite the current US administration's crackdown on relations with Cuba (US citizens are not allowed to visit as tourists) there are even some studying in the United States. At the University of Havana most professors are expected to spend a quarter of their time doing research (there are a few full-time researchers); research decisions are made at faculty council level but priority goes to areas that are being recommended by the Academy of Sciences. Again, applied biotechnology is very much to the fore.

There are informal collaborations with US scientists, particularly through the organization NACSEX (North American Cuban Scientific Exchange) run by Harlyn Halvorson of Brandeis University. But not as many as the Cubans would like. Dr Rudolfo Maribona of CENIC, who works on the molecular biology of eye spot disease in sugar cane, said plaintively "we invite the US scientists but they don't come". Be that as it may, Cuban biotechnology is making progress. **Tim Beardsley**

The Second Cuban Seminar on Interferons and First Cuban Seminar on Biotechnology, Havana, 20 – 22 February. Tim Beardsley's visit to Cuba was paid for by the Cuban National Academy of Sciences.

British civil plutonium

SIR—We are pleased that the Central Electricity Generating Board (CEGB) has responded (*Nature* 318, 406; 1985) to our paper (*Nature* 317, 213; 1985) in which we concluded that 2.3 ± 0.8 tonnes (te) of plutonium is missing from the UK civil stockpile. The calculations that we published last September resulted from a very thorough study of all the criticisms of our earlier work made by CEGB at the Sizewell Inquiry. We are pleased to see that only two main CEGB criticisms remain, and we will deal with both of them in turn. The first criticism is indicative of a rather superficial reading of our paper. The second criticism contains an argument that can easily be shown to be incorrect.

(1) It is definitely incorrect to say that our claims for ± 2 per cent accuracy were not substantiated. We presented a number of tests that showed our results to be within ± 2 per cent of published figures or, if not, to be underestimates. They included the following.

(i) We predicted that the total plutonium in the cores of the magnox reactors on 31 March 1985 was 9.6 te. Our calculations were completed before the government came up with its own figure of 9.50 ± 0.25 te. Here ours is a 1 per cent overestimate. We pointed out that our procedure slightly overestimates plutonium in core but not total plutonium production.

(ii) We predicted that plutonium stocks on 31 March 1983 and 31 March 1985 would differ by 4.9 te. Government figures (quoted to nearest 0.5 te) differ by 5.0 te. Here ours is a 2 per cent underestimate.

(iii) Since Sizewell, most of our effort has gone into deriving $G(B)$, plutonium isotope production in kg per te as a function of burn-up, from first principle incorporating all factors mentioned by CEGB in its non-quantitative criticisms. We particularly object to the description of our new station-by-station calculation of $G(B)$ as a "simplified method . . . which uses approximate data culled from a variety of sources". The reference list in our paper shows that the data come mainly from UK Atomic Energy Authority and International Atomic Energy Agency sources.

Among the tests of $G(B)$ we reported were: When using appropriate parameters, $G(B)$ exactly reproduces published data on plutonium production as a function of burn-up in the Calder Hall reactors (agreement well within 1 per cent); Table 1 in our paper demonstrates that using parameters appropriate to the civil reactors, we get good agreement with isotopic ratios in CEGB's despatch data (differences of about 2 per cent for four stations and less than 1 per cent for the other four; averages agree to better than

0.1 per cent); use of "worst case" parameters reduces the plutonium total by only 1.5 per cent.

(2) We are surprised that CEGB should repeat again the claim that Wylfa should be included when comparing production and despatch figures. The background to this is that the Inspector asked CEGB to provide the Sizewell Inquiry with data on plutonium production after 1977. CEGB did not comply with this request, but instead, provided figures on plutonium despatched from each station for the six fiscal years 1978–84. At seven of the eight stations, the spent fuel is stored in ponds before despatch to Sellafield. As the spent fuel cannot remain in the ponds indefinitely, we believe that, provided the period considered (6 years) is much greater than the time the elements remain in the ponds (approximately 1.2 years on average), then the total plutonium despatched should be similar to that produced in the 6-year period starting one year earlier.

Wylfa, however, has a dry-store where spent fuel can remain indefinitely and where plutonium is accumulating. It is totally inappropriate to include Wylfa in any such steady-state analysis.

If there was a backlog in all the ponds at the start of the 6-year period and this backlog was removed by the end of the period, then more plutonium would be despatched in 1978–84 than was actually produced in 1977–83. This is what CEGB

says happened, but it does not quantify the size of the backlog. Our calculations are consistent with this, because our production figures *underestimate* the plutonium despatched by 3.6 per cent. If CEGB's account of the UK/US plutonium exchange is to be correct, our figures have to be a 9 per cent *overestimate* of the actual figures.

There is an easy way for CEGB to discredit our calculations if they are so wrong. It could be done without risk to national security, as the 2 per cent error inherent in our calculations is equivalent to a large number of warheads. All CEGB needs to do is to publish the plutonium content of the spent fuel in the storage ponds at the start and end of the 6-year period or provide the information on plutonium production requested by the Inspector. CEGB refuses to do either, following advice from the Department of Energy, even though such data would help to resolve this controversy.

K.W.J. BARNHAM

Blackett Laboratory,
Imperial College,
London SW7 2BZ, UK

D. HART

School of Environmental Sciences,
University of East Anglia,
Norwich NR4 7TJ, UK

J. NELSON

London New Technology Network,
London NW1 9ES, UK

R.A. STEVENS

Intel Corporation, Santa Clara,
California 95051, USA

How big was the Ark?

SIR—In a reply to my earlier letter (*Nature* 316, 184; 1985) R.H. Clarke of Melbourne University (*Nature* 318, 100; 1985) chose also to ridicule the following statement of mine published in the Melbourne Age in 1984: "first, does he realize that Noah's Ark was immense ($450 \times 75 \times 45$ feet)? Its tonnage would have supported 4,000 fully grown African elephant bulls. There would have been plenty of room for all had God put young animals on board."

Clarke says that such beliefs should not be imposed on the young and vulnerable, but my statement is ludicrous only if taken out of context. Because of the tremendous genetic potential for variation within the species, the total number of representative ancestral pairs that had to be on board, to account for the diversity of extant terrestrial fauna, did not have to be as large as might be imagined. We need only to think of *Canis familiaris* (dogs) to realize that in theory a single pair on the Ark could have been sufficient. Invertebrates could have come on board via their plant or animal hosts. Interbreeding followed by natural selection (or artificial breeding) would have run its normal course in subsequent millennia.

Microevolution can proceed very rapidly, as has been demonstrated in experiments by the Commonwealth Scientific and Industrial Research Organization where five pure dingo pups mated with domestic dogs resulted in 41 hybrids possessing a wide range of characteristics which matched the diversity of wild dingo-dog hybrids now prevalent in the Australian bush. Since biblical chronology places the launch of the Ark at about 4,500 years ago, there would have been ample time to provide the diversity of species now extant. It is also relevant to my case that classification at the species level does not reflect breeding boundaries, as many cases of potential interspecific hybridizations within a genus have been documented (for example *Canis*, *Felis*, *Drosophila*). Thus, while microevolution (which includes the case of the peppered moth) is rapid and observable, direct evidence for macroevolution, which seeks millions of years for change, is still non-existent. My paradigm, to use Kuhn's terminology, might be different from that of macroevolution, but the facts need not be.

C.K. PALLAGHY

School of Biological Sciences,
La Trobe University,
Bundoora, Victoria 3083, Australia

Appeal from Moscow

SIR—I am appealing to you for help. For about seven years, I have been trying to leave the Soviet Union for Israel. All my adult life (more than 20 years) I have worked in the field of experimental oncology and I have never had anything to do with any state or military secrets. Nevertheless, for the past seven years, I have been refused permission to leave the country and have been prevented from working in my field.

I am appealing to all my friends, to all my colleagues, to "the world of oncology" for help and sympathy. I am convinced an appeal from you to the General Secretary M.S. Gorbachev to permit me to leave the country will be considered with due attention and I will be permitted to realize my legitimate wish.

Following the Geneva talks, relations are being restored between East and West and joint Soviet–American research projects are being set up in oncology; I therefore believe that such a request would be looked upon with understanding. A request from my US colleagues would be of particular value.

My own state of health and that of my mother make me anxious about any delay and I ask all of you "not to put off to do the good". You are my last hope. I believe in your solidarity and humanism.

IOSEF S. IRLIN

Moscow, USSR

Helping Sakharov

SIR—The letter by Erast B. Gliner¹, "Another reason for saving Sakharov", does more harm than good. By emphasizing the value of Sakharov's scientific contributions, Gliner reinforces the tendency of some scientists to protest against violations of human (and scholarly) rights in proportion to perceived eminence of the individual.

The history of science shows that eminence is often accorded only posthumously. Thus, such a criterion for intervention is likely to be too late, especially where intellectual suppression has prevented younger scientists from realizing anything approximating their full potential.

Also, there is a natural reticence among scientists to comment publicly on the quality of research that is outside their area of professional competence. Again, the history and sociology of science provide many examples of where individuals in the same field (who may be competitors, or may be protégés) have expressed bias²⁻⁵.

Thus, by using real or potential eminence as a criterion for protest, we greatly restrict the number of scholars who might be enlisted to the cause and, as Gerald Holton⁶ has shown, only a small fraction of the scientific community shows "societal concerns".

It is curious that it has taken so long for some scientists to learn the lesson of Amnesty International, which has had some relative success in dealing with unjust incarceration or torture in authoritarian regimes of the right, left and centre. There is a need for a strict apolitical approach, emphasizing deviation from common humanitarian ideals as expressed in the writings of both capitalist and socialist states. It is desirable to avoid unnecessary labelling or stereotyping, for example Gliner's¹ "dangerous 'bourgeois ideology'".

It is ironic that it was none other than Sakharov himself who complained that the unfair political slant of an earlier item in *Nature* "almost cancelled its usefulness"⁷.

CLYDE MANWELL
C.M. ANN BAKER

Department of Zoology,
University of Adelaide,
South Australia, 5001

1. Gliner, E.B. *Nature* **318**, 513 (1985).
2. Council for Science and Society *Scholarly Freedom and Human Rights* (Barry Rose, London, 1977).
3. Martin, B. *The Bias of Science* (Society for Social Responsibility in Science, Canberra, Australia, 1979).
4. Manwell, C. & Baker, C.M.A. *Search* **12**, 151–160 (1981).
5. Peters, D.B. & Ceci, S.J. in *Peer Commentary on Peer Review* (ed. Harnad, S.) 3–11 (Cambridge University Press, 1982).
6. Holton, G. *The Scientific Imagination Case Studies*, especially pp. 234–235 (Cambridge University Press, 1978).
7. Sakharov, A. *Nature* **288**, 112 (1980).

Evolution again

SIR—Attacks on and defence of Charles Darwin and his evolutionary views are obviously unavoidable as a result of the current debate about the biological evolution. Dover recently emphasized the importance of the phenomenon of *stasis* in this context¹. I, however, would like to suggest that Darwin, for psychological and strategic reasons, of which he was perhaps not quite aware, excluded, *a priori*, *stasis* as well as evolutionary jumps from his definite theory of transmutation. Although in retrospect Darwin's contemporaries might appear to have been prepared for the shift of epistems, the successful introduction of the idea of transmutation into the scientific community was a huge task for Darwin; it was even held to be responsible for his mysterious invalidism². From his conviction of *permanent* and *gradual* change in nature, Darwin eventually got the required intellectual energy to accomplish his project. The orthodoxy, as is well known, advocated the fixity of species and acts of special creation — the first can be considered as an extreme form of *stasis*, the latter of jumps. Whereas Darwin himself hinted at the phenomenological similarity between special creation and evolutionary jumps³, the implication of this similarity between *stasis* and the fixity of species should also be considered in the discussion of the development of his theory. Darwin's strategy then, was to propose an opposite

picture of the living world rather than explain the old issues in terms of transmutation. If he had incorporated a plausible explanation of *stasis* into his theory, the orthodoxy would probably have used it as an argument in favour of the fixity of species.

This approach was counteracted by Darwin's obsessional reverence for facts, whether they suited him or not. Consequently in *The Origin of Species* Darwin included significant tenets of the new hypothesis of punctuated equilibrium as was shown by Rhodes⁴.

MIRKO MAJER

Mozartstrasse 5,
D-8510 Fürth/Bayern, FRG

1. Dover, G.A. *Nature* **318**, 19 (1985).
2. Colp, R. *To be an Invalid* (Chicago and London, 1977).
3. Darwin, C. *The Origin of Species by Means of Natural Selection* 6th edn, 424 (Murray, London, 1880).
4. Rhodes, F.H.T. *Nature* **305**, 269 (1983).

Classical relativity

SIR—H. Aspden (*Nature* **318**, 317; 1985) says: "Relativity, in its classical sense, will undoubtedly survive, but the real question is whether the test of time will continue to support Einstein's requirement that each and every relatively-moving observer constitutes his own personal reference frame for the locally-applicable laws of physics". Everyone would agree with the latter suggestion, but what is the meaning of the statement "Relativity, in its classical sense, will undoubtedly survive?"

N.J. KOLYVODIAKOS

"Studies in Physics",
12 William King Street,
GR-11253 Athens,
Greece

H. ASPDEN REPLIES—N.J. Kolyvodiakos has questioned the meaning of "Relativity, in its classical sense". This was a reference to the pre-Einstein version of relativity. This is that physical phenomena measured exclusively within the confines of a system in non-accelerated motion will not reveal that motion to an observer transported with the system.

Einstein went beyond this by relying on the additional axiom that light propagates in free space with the same speed relative to all observers. This axiom was not restricted to light speed in the immediate vicinity of the observer or to observers sharing the motion of the apparatus by which light speed was measured. The observers could be in relative motion, even though close together and viewing the same ray of light, however remote it might be. This has given scope for paradox and, though most of the theoretical criticism has been parried, it encourages the search for experimental violation of Einstein's version of relativity.

Department of Electrical Engineering,
University of Southampton,
Southampton SO9 5NH, UK

Proprietary rights to research

The rules that say when authors should make research materials available to others cannot be hard and fast. But release should be the rule, not the exception.

To what extent are researchers who have announced some discovery of importance beholden to colleagues (and competitors) in other laboratories to provide research materials necessary for the replication of the work reported? This is the question vividly raised by the exchange of letters appearing on pages 21 and 22 of this issue. The issues raised are ancient, common and of a personal character.

The circumstances are these. Two years ago, Gusella and his associates described the isolation of a genetic probe for a part of the human chromosome 4 lying so close to the unidentified gene responsible for Huntington's chorea that it seemed as if it might be a diagnostic agent for the disease. Two years ago, the technique by which this had been done was relatively new. It rests on the facility for chopping chromosomal DNA into pieces of different lengths by the use of restriction enzymes. Gusella's finding of a putative genetic probe for Huntington's chorea was one of the first successes of this kind; in the interval, specific probes for other genetic diseases (Duchenne muscular dystrophy and cystic fibrosis, for example) have come to light.

The nature of the technique is crucial to an understanding of the dispute between Gusella and a group in Oxford, which has been itching to get its hands on the Huntington's probe during the past two years. Not all the DNA in the human (or any other genome) is meaningful. Some consists of sequences of nucleic acids that encode essential protein molecules, but some of it is a "nonsense" DNA in the extreme sense; it may consist only of mindless stretches of apparently meaningless repeating units. Between these extremes, there are stretches of DNA whose placing and character play essential parts in the regulation of gene expression. Still others appear to have a function no more specific than that of holding together essential pieces of DNA at the correct spacing, and can thus continue to function even if quite substantial variations of the nucleotide sequence occur.

Because restriction enzymes cut DNA molecules only at places where the nucleotide sequence is that specific to their own function, variations among the genome of individuals will, now and again, yield places where the genome may, or may not, be cut by a particular enzyme. The upshot may be a variation between one person's genome and another's, revealed

by the occurrence of differently sized fragments. Experience in the past few years has indeed shown that these differences are inheritable. This polymorphism is exactly analogous to that in which functional proteins such as the globin of haemoglobin may vary from one person to another and, given parallel inheritance from two parents, may consist of a mixture of the two forms.

What Gusella *et al.* described two years ago was the identification of the site of a restriction fragment length polymorphism, somewhere on chromosome 4 in human beings, which appears to be closely linked in its inheritance with the gene responsible for Huntington's chorea, whatever that may be. But the polymorphism and the gene are some way apart from each other. Gusella and his colleagues were at pains to insist, in their account of their work, that their "probe", a piece of DNA spanning the variable restriction site, should not, for the time being, be used as a diagnostic test for the disease. Plainly, there is a chance that the restriction site and the gene will not be inseparably tied together in inheritance, and until there are probes for sites closer to, or on either side of, the putative gene, there will always be a possibility of false diagnosis. There is, moreover, still the possibility that some families with Huntington's disease will prove to have a completely different genetic defect.

This is the basis on which Gusella has so far declined to provide the structure of his group's probe to the Oxford group, which says that it wishes to "apply DNA techniques to the practical problems presented by patients and their relatives". The need is easily imagined. Huntington's chorea is perhaps the most haunting of inherited diseases. Throughout early life and especially during many procreative years, patients will be free of symptoms. But then they are afflicted with neurological deterioration which is invariably fatal. The Oxford group is bent on genetic counselling among the relatives of cryptic victims of the disease, which would entail that many unknowing future patients should be told of their condition. Gusella says that this will remain unethical so long as the diagnosis is less than certain. He is concerned above all with the plight of the false positives, those who may carry a restriction fragment length polymorphism characteristic (for the family concerned) of Huntington's chorea.

On this narrow issue, Gusella may be over-squeamish. People belonging to families in which Huntington's is rife are probably already painfully aware of the risks to themselves and their children. These may well be the circumstances in which firmer knowledge, however much more painful, is preferably to gnawing anxiety. But one such as Gusella, who has at his disposal a material that may be of value in the diagnosis of a genetic disease, would be well within his rights to require a full explanation of the programme of investigations that others wish to undertake with his material. Indeed, it would be irresponsible to do otherwise.

This points to a more general issue, that of the obligations of researchers to others wishing to obtain samples of research material. It is well within the rights of a researcher in such a position to demand a full account of the purposes for which the material is required. Mere curiosity is an insufficient basis for a request. The possibility that research material might be used for commercial purposes is also a sufficient reason why a researcher might decline to provide it on demand. This shows that the obligation of a researcher to provide research materials cannot be absolute.

There are other circumstances in which refusal cannot be justified. Researchers have a legitimate right to verify what their colleagues and even competitors claim to have done. Artificial restrictions, such as the suggestion that verification should be carried out in the original laboratory, are only occasionally justifiable (as on safety grounds). Even allowing that the provision of research materials may help a competitor to get ahead should not be grounds for refusal. The principle that discoveries are common property is widely acknowledged but not always gracefully accepted.

Nature has not so far followed some other journals in making an undertaking to supply research materials a condition of publication because of difficulties such as that to which Gusella refers. But its interest, like that of the scientific community as a whole, lies in that direction. A promise that research materials will be made available to others will not be a firm condition of publication, but the presumption that reasonable requests will be met will remain an ingredient of polite correspondence with would-be authors.

John Maddox

Astrophysics

Discovering a bubbly Universe

from Joseph Silk

PERHAPS the most spectacular of the large voids in the three-dimensional distribution of galaxies is the Bootes void, a region at least 50 megaparsecs in diameter that contains no luminous galaxies. Until now such voids have seemed interesting oddities that are largely irrelevant to the large-scale structure of the Universe and to the clusters and superclusters of galaxies. But a new survey of the large-scale galaxy distribution by Margaret Geller, John Huchra and Valerie de Lapparent, published in the 1 March issue of *Astrophys. J. Lett.*, reveals that large voids are not the exception but the rule. The implications for our understanding of the origin of the large-scale structure are profound.

The survey is a systematic collection of redshifts of all galaxies of apparent magnitude brighter than 15.5 in a region measuring 6 degrees by 117 degrees on the sky. The redshifts, via Hubble's law, provide a measure of distance from the observer, giving a three-dimensional map of the galaxy distribution in a limited volume of the Universe. Inspection of the map (see figure) reveals a striking result: large, apparently empty, quasi-spherical voids dominate space, and the galaxies are crammed into thin sheets and ridges between the holes. Similar maps have been presented previously by Jaan Einasto and colleagues, but the reality of this phenomenon was largely discounted because of the inhomogeneity of their data set. The new data thoroughly sample a localized region of the sky, and unambiguously demonstrate the foam-like appearance of the galaxy distribution.

Two notes of caution deserve to be sounded. No correction is made for galaxy-peculiar velocities; these are motions relative to the Hubble flow that are especially large in rich galaxy clusters and will tend to distort the three-dimensional maps that naively translate redshift into distance. This produces the 'fingers-of-God' effect: elongations of the galaxy distribution towards the observer that are not really present appear in the redshift dimension. Another difficulty is that in any magnitude-limited survey, a broad range of nearby galaxies is sampled. However, the most distant galaxies in the sample are exclusively the most luminous ones. If galaxies of different luminosities vary in their spatial distribution, this could also lead to a systematic distortion of the derived maps; for example, the most luminous galaxies, like the highest mountain peaks, may be more clustered than galaxies of average luminosity.

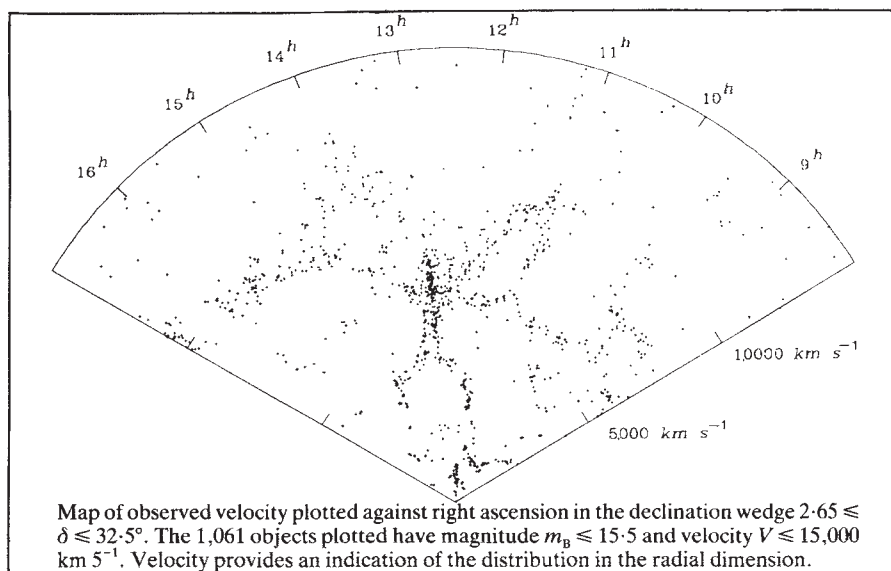
The largest bubbles are 30–50 mega-

parsecs in diameter. Of note are the well-defined edges of the galaxy layers. How could this situation arise? There are two possibilities: either galaxies formed selectively, at only the edges of the voids; or they moved after formation so as to evacuate large voids.

Consider the former case. A theory that yields beautiful voids was developed by

form the voids. A more serious difficulty is that even the most optimistic estimates of energy output fail to create sufficiently large voids. Something else is needed.

The missing ingredient is gravity. Density fluctuations were present at the beginning of time, in the earliest instants of the 'big bang', and gravity amplifies these fluctuations into the large-scale structure of the Universe. Most cosmologists believe that galaxies originated in this manner rather than by explosive amplifications of primeval seeds, which themselves must be attributed to initial conditions. The large-scale structures (notably the clusters and superclusters) and the voids



Jeremiah Ostriker and Lennox Cowie (*Astrophys. J. Lett.* **243**, 127; 1981), and independently by Satoru Ikeuchi (*Publ. astr. Soc. Japan* **33**, 211; 1981). According to Ostriker and Cowie, an explosion initiated by many supernovae in a newly forming galaxy drives a blast wave that propagates outwards and sweeps up a spherical shell of ambient gas. A hole is evacuated, and the unstable compressed shell fragments to form more galaxies. These, in turn, develop blast waves; a series of bubbles develop that fill most of space, with galaxies present only on shell surfaces.

Explosive galaxy formation is a seductive hypothesis that reproduces the observed foam-like nature of the galaxy distribution. But there are two difficulties that make most cosmologists sceptical of its applicability. One is the plausibility of the mechanism itself — supernovae explode and clear out holes that are tens or, in rare cases, hundreds of parsecs across, but does this phenomenon really work on scales of tens of megaparsecs? Billions of supernovae are presumed to explode coherently, that is to say, over the crossing time of a galaxy of about 10^8 years to yield a vast explosion. A gas-rich protogalaxy could conceivably have quenched such explosions by radiating away much of the kinetic energy of bulk motion needed to

must certainly have formed by the action of gravity in acting over tens or hundreds of megaparsecs. The advantage of this approach is that the physics of a system in which the only significant force is gravity is well understood. It is necessary to use a large computer to perform a numerical simulation, but statistical measures of large-scale clustering, such as the galaxy correlation functions, are known (see Davis, M., Frenk, C., Efstathiou, G. & White, S.D.M. *Astrophys. J.* **292**, 371; 1985). But to understand the great voids and the sheet-like regions occupied by galaxies it is necessary to go one step further. The concept of biasing the formation of large-scale structure was introduced by Nick Kaiser and has recently been applied to galaxy formation (Peacock, J.A. & Heavens, A.F. *Mon. Not. R. astr. Soc.* **217**, 805; 1985, and Barden, J., Bond, J.R., Kaiser, N. & Szalay, A. *Astrophys. J.* in the press). Galaxies are presumed only to form in the rare peaks of an initially gaussian distribution of density fluctuations. Peaks that occur in a potential large-scale cluster acquire a slight boost that enables galaxies to form; those elsewhere fall below the threshold. The biasing hypothesis enhances the large-scale structure that develops as gravitational forces amplify the initial fluctuations.

Voids develop, containing hundreds of 'failed' galaxies, and galaxies form only in the dense regions between voids.

Biasing enables simulations of a universe containing cold dark matter at the critical density for closure to be reconciled with the observational determinations of the density parameter of the Universe. The voids are not really void, but contain matter that has failed to become luminous. The dark matter is more uniformly distributed than the luminous matter, and is an inert background that does not respond to most astronomical tests. The dark matter is weakly interacting and clusters on all scales (hence the label cold). It selectively forms galaxies at an early epoch in the rare density peaks. The theory is deficient in two respects: it fails to reproduce the most massive aggregates of galaxies and to account for the clustering of the great galaxy clusters; and the biasing hypothesis lacks any convincing physical explanation.

It could be that neither the explosive amplification scenario nor the biased cold dark matter scenario will account for the foam-like distribution of galaxies that is emerging from the deep surveys. This is no reason for despair; one possibility, for example, is that a scale of 50 megaparsecs or so could have been built into the initial conditions of the Universe. This is not quite as implausible as it seems — in fact, speculations centred around 'hot' dark matter can readily incorporate a scale of this magnitude. Hot dark matter, usually associated with the massive neutrino, has fallen into disfavour because there is too little small-scale power in the initial fluctuations to explain simply how galaxies formed. But the large-scale distribution of matter seems to fall out naturally, rather like a tidal wave sweeping all before it.

One class of theories forms galaxies but has difficulty accounting for the large-scale galaxy distribution, and another class explains large-scale structure but cannot simply form galaxies. Yet another approach is to abandon the assumption that the primordial fluctuations were part of a gaussian distribution. With non-gaussian fluctuations, rare peaks may stand out, whereas the average level of density fluctuations still remains low to conform with the observed isotropy of the cosmic background radiation. The best-studied realization of such a theory invokes vacuum strings left over from a phase transition in the very early Universe. These topological defects create sites around which galaxies accrete, and can also account for the large-scale clustering of the galaxies. It remains to be seen whether such exotic solutions are necessary to meet the challenge of the new redshift survey. □

Joseph Silk is in the Department of Astronomy, University of California, Berkeley, California 94720, USA.

Species conservation

The cautionary tale of the black-footed ferret

from Robert M. May

IN September 1981, a colony of black-footed ferrets (*Mustela nigripes*), a species thought by many to be extinct, was discovered near Meeteetse, Wyoming, in the United States. The numbers of these animals in the wild grew steadily from 60 in 1982 to 128 in 1984 (ref. 1). But this trend came to an abrupt end in the summer of 1985 when the ferrets became infected with canine distemper. By mid-January of this year, only three individuals are known to remain in the wild and six are alive in captivity. The future of the species could depend on these six animals, and the history of events can be read as a cautionary tale.

The black-footed ferret, the North American representative of a Holarctic group of polecat species, is closely related to the European ferret, which was used for hunting in earlier times and is now domesticated as a pet in Europe and North America. The main prey of the black-footed ferret is the prairie dog, a burrowing member of the squirrel family, which used to be abundant (in large complexes of 'towns') on the North American prairies.

In the first half of this century, extensive campaigns of prairie-dog poisoning accompanied agricultural development, and the abundance of these animals dropped by at least an order of magnitude, together with a contraction of their range, which once stretched from the Rocky Mountains to the Great Plains, and from Mexico to Canada. The numbers of black-footed ferrets declined precipitously in parallel with the decrease in numbers of prairie dogs. Rising concern about the disappearance of the native prairie ecosystem led to the first systematic research on black-footed ferrets in 1964, but by 1974 so few animals remained that the programme was discontinued, and by 1980 many believed the ferret to be extinct¹.

In 1981, a wild ferret was killed by a domestic dog near Meeteetse, and the realization that the animal was not extinct led to a study of the colony by researchers from the US Fish and Wildlife Service and from the Biota Research and Consulting company. Tracking ferrets is made difficult by their nocturnal, underground lifestyle; numbers are estimated each summer by spotlighting (repeatedly counting the number seen using high-intensity light flashes bit-by-bit throughout the 80,000-acre colony) and by mark-recapture (trapping and marking about one-third of the population twice each summer)². The

numbers thus estimated have been more than 60 for the summer of 1982, 88 in 1983 and 128 in 1984.

This apparent pattern of population growth came to an abrupt end in the summer of 1985. The first count in August gave 58 ferrets. By October the number (estimated by mark-recapture) was 31. A few days later, the cause of the decline was identified by Tom Thorne, the veterina-

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

WWF/Helmut

Black-footed ferret

rian of the Wyoming Game and Fish Department: the ferrets were infected with canine distemper, which is fatal. After some delay, 12 ferrets were brought into captivity, 6 of which have since died of distemper. It is hoped that the remaining 6 (2 young males, 4 females) are beyond the incubation period for the disease.

This tale prompts many questions. Why did the numbers of ferrets decline more dramatically than the abundance of their prairie-dog prey earlier in the century? Why did the one known population of ferrets at Meeteetse exhibit apparent growth from 1982 to 1984, or was this a minor fluctuation? Is the outbreak of canine distemper associated with the population's 'miniboom', or with the intrusive presence of the researchers?

Some would answer the first question by observing that predator populations often exhibit disproportionate crashes when their prey populations diminish. The reasons for this are not agreed. My own view is that relatively specialized prey-predator associations typically require spatial patchiness for their persistence: the associations are often a shifting mosaic of empty patches; prey-only patches where prey populations are doing well; and patches with prey and predator where predators are currently flourishing. The dynamics of such systems obviously depend on the rates of prey and predator movement among patches and on the demographic details of prey-predator interactions within any one patch, but they can also depend in a complicated and non-linear way on the number of patches and on the overall prey density. Halving the average prey density, or halving the num-

ber of patches (or the geographical range of the prey), can produce qualitative changes³.

These generalizations do not provide a detailed explanation for the fluctuations of the ferret population at Meeteese. But it could be that the population had been steadily increasing in the early 1970s until it exceeded the threshold density that allowed an epidemic of canine distemper to be propagated once the infection was introduced by a raccoon, coyote or other carrier. In this view, the events at Meeteese — although seen in isolation today — may be one typical element from the shifting mosaic that prevailed in the past. Such a situation is not without parallel elsewhere in the carnivore order: although their social organization is very different from ferrets, silver-backed jackal populations of the Serengeti appeared to have exhibited an increase in average density (and a contraction in average territory size) in the past few years, followed very recently by a crash that may be associated with an infectious disease (probably canine distemper)⁴. Wild dogs on the Serengeti are also known to have suffered from periodic outbreaks of disease during the past 30 years (again, probably distemper)⁵.

The situation in Meeteese was made more complicated by the discovery that sylvatic plague has been present among the prairie dogs since June 1985. The decrease in numbers of prairie dogs never seemed sufficient to account for the decline in ferrets (although there was argument about this point until the canine distemper was identified), but this decrease in the food base can be argued to have reduced the resistance of the ferrets to infection. Note that prairie dogs are thought not to be affected by canine distemper, nor are ferrets by sylvatic plague.

The notion that parasites and pathogens often affect the abundance or geographical distribution of natural populations is resisted by many wildlife biologists^{6,7}. In part, this resistance is grounded on the belief that host-parasite co-evolution will tend to produce parasites that are relatively harmless to individual hosts, and which therefore have little effect on the dynamics or biogeography of the host population⁸. Conspicuous exceptions to this rule — rinderpest in East Africa or the infections that create problems among laboratory or zoo populations — are dismissed as resulting from exposure of the host populations to unfamiliar pathogens, or (more tortuously) familiar pathogens in an unfamiliar setting.

More accurately, evolutionary pressures act on individual parasites to maximize reproductive success. This may involve complicated trade-offs between transmissibility and virulence — that is, between producing transmission stages to infect new hosts and harming the host in

the process. The endpoint is not necessarily a harmless pathogen, as witness the many lethal viruses, bacteria and protozoans that afflict invertebrate species, or the infections that modify host behaviour in deleterious ways to enhance their own transmissibility. Evolution produces what is best for individual parasites, which is not always what is best for the parasite population⁹.

Whatever the reason, viral, bacterial, protozoan and helminth parasites are rarely mentioned in ecology texts or wildlife management courses as factors that can influence the abundance, or even the survival, of host populations.

These poorly understood processes provide the background for the events at Meeteese. For the past two years, various scientists and conservation organizations (including the World Wildlife Foundation and the New York Zoological Society, who have supported much of the research) have argued that a programme of captive breeding should be established. These arguments became insistent, and to my mind quite unanswerable, when the ferret population was known to be more than 100 in the summer of 1984. But the head of the Wyoming Game and Fish Department's ferret programme, Harry Harju, remained intransigently opposed to any capture of ferrets for such a breeding programme. Harju's reasons for passively letting events take their course are difficult to fathom — he appears to have been often more concerned with bureaucratic issues than with biological ones. Wyoming officials are reported to have been determined to keep all the ferrets in Wyoming, because otherwise (in Harju's words) "We'd have no control over them". Harju is also reported¹ as saying "No one, but no one, suggested taking ferrets for nonsensical reasons. If you really look at it, they

all had selfish motives". Although this may be true, it may reveal more about the accuser than the accused.

If such a mess can be made of efforts to save an attractive creature such as the black-footed ferret in a country as well organized and prosperous as the United States, prospects for conservation in other parts of the world are indeed bleak. I do not doubt that Harju will in time attain mythical status in the conservationist's pantheon as a symbol of what can happen when the machinery and dignity of office come to overshadow the purposes the office was created to serve.

The morals drawn here will be recognized, with a greater or lesser degree of enthusiasm, by most protagonists in the Meeteese ferret saga. I conclude by emphasizing a different moral. Diseases are important in many natural populations, particularly in disturbed situations, and I advise special concern about the possibility of disease in the case of a species that is rare and yet whose population is locally undergoing steady growth. □

1. Weinberg, D. *Nat. Hist.* 95(2), 62 (1986).
2. Richardson, D. *Nat. Hist.* 95(2), 68 (1986).
3. May, R.M. & Beddington, J.R. in *Problems in Management of Locally Abundant Wild Mammals* (eds Jewell, P.A. & Holt, S.) (Academic, New York, 1981).
4. Mochman, P.D. *Adv. Mammal Behav. Spec. Pub.* 7 423 (1983).
5. Sinclair, A.R.E. in *Serengeti: Dynamics of an Ecosystem* (eds Sinclair, A.R.E. & Norton-Griffiths, M.) 13 (Chicago University Press, 1979).
6. Hassell, M.P. *et al.* in *Population Biology of Infectious Diseases* (eds Anderson, R.M. & May, R.M.) 15 (Springer, New York, 1982).
7. Holmes, J.C. in *Population Biology of Infectious Diseases* (eds Anderson, R.M. & May, R.M.) 37 (Springer, New York, 1982).
8. Allison, A.C. in *Population Biology of Infectious Diseases* (eds Anderson, R.M. & May, R.M.) 245 (Springer, New York, 1982).
9. May, R.M. & Anderson, R.M. in *Coevolution* (eds Futuyma, D.J. & Slatkin, M.) 186 (Sinauer, Sunderland, Massachusetts, 1983).

Robert M. May is Class of 1877 Professor of Zoology at Princeton University, Princeton, New Jersey 08544, USA.

Bacterial chromatin

A new twist to an old story

from David Lilley

DESPITE their relative simplicity, we know much less about the way bacteria organize their DNA structurally than we do about the chromatin of higher organisms. But this situation may be about to change. Chromosomal DNA purified from either source is supercoiled to about the same extent. The DNA is unwound (negatively supercoiled) by about 1 Watson-Crick turn in every 20, that is, it has a superhelix density of -0.05 .

Supercoiling a closed circular DNA molecule¹ results in structural deformation, which can be of two kinds (see figure). The helical period can increase by deformation about the duplex axis (twist), or the duplex axis itself may deform in

space (writhe). The algebraic sum of these two deformations is a constant, but in a pure DNA molecule there is free partition between twisting and writhing². Here, the DNA molecule is torsionally stressed, providing energy which can be coupled to the opening of the double helix, or the formation of local structural variants such as left-handed Z-DNA or cruciforms.

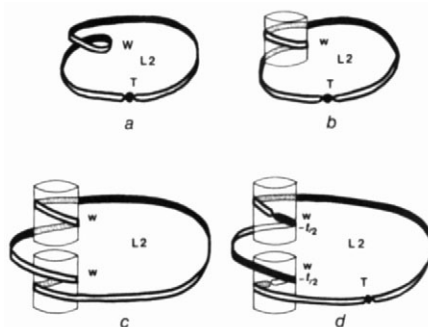
Recent experiments suggest that the effective superhelix density inside the living bacterium is considerably lower than the -0.05 of extracted DNA. The structural transitions in DNA can be used to measure local superhelix density. Peck and Wang³ noted that the B-Z transition of a purine-pyrimidine stretch (C-G)₁₀,

which occurs at a threshold superhelix density of -0.04 , prevents the passage of a transcribing RNA polymerase ternary complex *in vitro*, whereas the same $(C-G)_{16}$ sequence does not impede transcription *in vivo*. Presumably the sequence is unable to undergo the B-Z transition *in vivo* because the effective superhelix density is lower than the -0.04 required for thermodynamic stability of the Z conformation. Experiments using the B-cruciform transition of an $(A-T)_{34}$ stretch suggest that the superhelix density is even lower *in vivo*. Greaves *et al.*⁴ showed that the $(A-T)_{34}$ sequence forms a cruciform structure *in vitro* at a superhelix density of -0.027 , and unlike other cruciforms, has no kinetic barrier to hinder the interconversion. But, like the $(C-G)_{16}$ stretch, this sequence does not seem to interconvert to its alternative structural form inside the bacterium, implying that the effective superhelix density is less than -0.027 .

The behaviour of a range of $(A-T)_n$ sequences in bacteria allows a fairly precise estimate of the level of superhelix density inside the cell. When n is 34 or less the sequences remain genetically stable over many generations. In contrast, an $(A-T)_{37}$ sequence is fairly unstable to deletion (R. Patient, personal communication), a behaviour typical of long inverted repeats^{5,6}. It is probable that the extra length of the $(A-T)_{37}$ sequence, that is, the larger change in local twist, places this cruciform at the point at which it is of borderline stability at *in vivo* superhelix densities. Assuming that the free energies of cruciform formation of $(A-T)_{34}$ and $(A-T)_{37}$ are similar, the superhelix density at which the latter just becomes energetically stable is calculated to be -0.024 .

J. Bliska and N.R. Cozzarelli (personal communication) looked at site-specific recombination in phage λ integration. They observed that, *in vitro*, the products depend on the superhelix density of the parent molecule in a linear manner. Thus, by examining the products of the reaction *in vivo*, they estimated the level of superhelix density inside the bacterium to be -0.025 . The values obtained by the two independent approaches are in remarkably good agreement. They suggest that the *in vivo* supercoiling of bacterial DNA is about half of that previously supposed, and because the energy of supercoiling is related to the superhelix density quadratically, the supercoiling energy available is only one quarter of the *in vitro* level.

Why is the level of supercoiling lower inside the cell? The effect is not caused by the action of topoisomerases; the linking number (the fundamental topological property which measures the linkage between opposite DNA strands) is unchanged. In eukaryotes, although the DNA has a superhelix density of -0.05 when purified, chromatin is largely relaxed because of the binding of histones to form nucleo-



Distribution of twisting and writhing in a ribbon (DNA molecule) with two superhelical turns ($L2$). *a*, Two turns free of constraint, shown here as one writhing turn (W) and one twist (T). *b*, Equivalent to *a* except that one writhing turn has been fixed by wrapping around a protein former (w). There is now only one turn remaining which is unconstrained, and thus the effective superhelix density has been halved. *c*, Two protein formers have now been placed on the ribbon ($2 \times w$). Two writhing turns are therefore fixed, leaving no superhelical turns freely available in the DNA. This DNA is effectively relaxed, that is, torsionally unstressed. This is equivalent to the state of the bulk of eukaryotic chromatin. *d*, Equivalent to *c* except that in addition to fixing a writhing turn (w), the protein core results in untwisting the ribbon by half a turn ($-t/2$). The net effect is that each protein bound only relaxes half a superhelical turn, thus leaving the DNA with half of the potential superhelix density, despite condensing the DNA to the same extent as in *c*.

somes. In the nucleosome, the DNA describes a left-handed toroidal helix which is fixed by a protein core⁷. Thus, the inherent supercoiling is completely accounted for by permanent writhing in nucleosomes, leaving no partition to twist and no net torsional stress in the DNA, except for a possible sub-population of active chromatin^{8,9}. Could protein binding account for the lowering of the bacterial superhelix density? Structural DNA-binding proteins, the HU proteins^{10,11}, are present in stoichiometric quantities and may have a histone-like role¹².

The involvement of protein is suggested when antibiotics are used to prevent protein synthesis in bacteria. Although B-Z and B-cruciform transitions are not normally found in bacteria, they can be induced by the addition of chloramphenicol to cells^{13,14}. Under these conditions the reduced availability of DNA-binding protein may cause the effective superhelix density to rise until the new structural variant becomes thermodynamically stable.

Broyles and Pettijohn¹⁵ provide a satisfying explanation for all these data and open up a new phase in the study of the bacterial chromosome by looking at the complex formed between HU proteins and DNA. The usual problem in such studies is in obtaining data on how linkage is partitioned between twist and writhe, as these parameters are frequently undetermined and hence cannot be unam-

biguously decomposed. The helical periodicity of the DNA in the complex was determined by high-resolution nuclease-cleavage studies, showing that cuts were introduced into each strand with a periodicity of 8.5 base pairs compared with the value of 10.5 seen for normal DNA¹⁶. These results suggest that the DNA becomes more tightly wound when it is complexed by HU, its periodicity being reduced from 10.5 to 8.5 base pairs per helical turn. The linkage change can be calculated by standard topological measurements on circular DNA molecules with bound HU; as the twist is known from the nuclease studies the writhe can be calculated. Like the nucleosome, a left-handed toroid is formed, with one turn of superhelix to three or four HU monomers.

The structure which emerges is similar to eukaryotic chromatin, but with one major distinction. The HU complex restrains the same number of superhelical turns as the nucleosome (it compacts the DNA to a similar extent) but in so doing fails to relax all of the torsional stress. The writhing change is partially reversed by an unprecedented overwinding of the duplex with the net effect that the DNA can be condensed while retaining a level of supercoiling about half that of the protein-free DNA. This agrees well with estimations of the effective superhelix density *in vivo*.

Why should bacteria evolve a DNA-binding protein with novel characteristics in order to retain a degree of supercoiling? Supercoiling increases the structural repertoire of DNA, facilitates strand opening and can influence promoter strength in both positive and negative directions. Despite the need to condense the chromosomal volume, the bacterial chromosome is divided into topologically independent domains and there is therefore considerable scope for control of gene expression at this level. □

1. Vinograd, J., Lebowitz, J., Radloff, R., Watson R. & Laipis, P. *Proc. natn. Acad. Sci. U.S.A.* **53**, 2204 (1965).
2. Fuller, F.B. *Proc. natn. Acad. Sci. U.S.A.* **68**, 815 (1971).
3. Peck, L.J. & Wang, J.C. *Cell* **40**, 129 (1985).
4. Greaves, D.R., Patient, R.K. & Lilley, D.M.J. *J. molec. Biol.* **185**, 461 (1985).
5. Collins, J. *Cold Spring Harb. Symp. quant. Biol.* **45**, 409 (1981).
6. Lilley, D.M.J. *Nature* **292**, 380 (1981).
7. Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D. & Klug, A. *Nature* **311**, 532 (1984).
8. Lilley, D.M.J. *Nature* **305**, 276 (1983).
9. Ryoji, M. & Worcel, A. *Cell* **37**, 21 (1984).
10. Rouviere-Yaniv, J., Gros, F., Haselkorn, R. & Reiss, C. in *The Organization and Expression of the Eukaryotic Genome* (eds Bradbury, E.M. & Javaherian, K.) 211 (Academic, New York, 1977).
11. Varshavsky, A.J., Nedospasov, S.A., Bakayev, U.V. & Georgiev, G.P. *Nucleic Acids Res.* **4**, 2725 (1977).
12. Rouviere-Yaniv, J., Yaniv, M. & Germond, J.-F. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3428 (1979).
13. Haniford, D.B. & Pulleyblank, D.E. *Nature* **302**, 632 (1983).
14. Haniford, D.B. & Pulleyblank, D.E. *Nucleic Acids Res.* **13**, 4343 (1985).
15. Broyles, S.S. & Pettijohn, D.E. *J. molec. Biol.* **187**, 47 (1986).
16. Rhodes, D. & Klug, A. *Nature* **286**, 573 (1980).

David Lilley is a Reader in Biochemistry and a Royal Society research fellow at The University, Dundee DD1 4HN, UK.

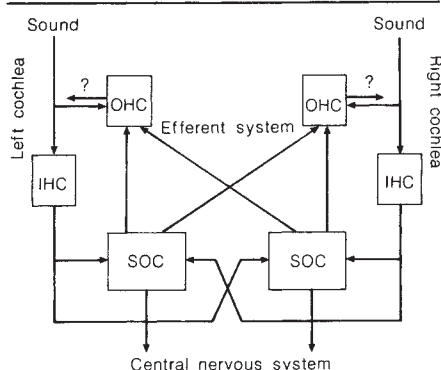
Cochlear neurophysiology

In one ear and out the other

from Jonathan Ashmore

Most people can give a reasonably coherent account of how the eye works. The same is probably not true of the ear, although some degree of hearing loss is the ultimate fate of everyone. One of the difficulties is that the cochlea, the hearing organ in the mammalian inner ear, is inaccessible and its structure complex. Another difficulty is that describing sound transduction in the cochlea depends on an understanding of mechanical events at the cellular and the molecular levels, and the techniques that can provide this type of information are relatively recent. Nevertheless, deafness arising from damage to any one of the cochlear elements is usually irreversible, and surgical remedies involving direct stimulation of the auditory nerve by electrodes implanted in the cochlea are being pursued energetically¹. The paper by Dodson *et al.* on page 65 of this issue² suggests a note of caution, but at the same time throws a curious light on one of the more enigmatic structures of the ear, the cochlear efferent system.

The inner ear in mammals separates component tones of a sound as distinct patterns of vibration along the basilar membrane by a mechanism essentially described by von Békésy more than 50 years ago, but now known to require some modification. In man, the basilar membrane is about 34 mm long and, for packing rather than for functional reasons, is coiled within the temporal bone. Movement of each section of the basilar membrane is sensed and encoded in the auditory nerve by a lever system involving two classes of sensory cell, the inner and the outer hair cells.



Relation of the efferent system to the inner (IHC) and outer (OHC) hair cells and to the superior olivary complex (SOC). The efferent system contains axons originating bilaterally in the brainstem. The mechanism by which OHC influence IHC is unknown. For simplicity additional features, such as relay nuclei of the auditory nerve and projections to the IHC have not been included (see ref.9 for details).

It is generally agreed that the inner hair cell is the primary sensory cell. But what are the outer hair cells doing? Hearing losses are often associated with the destruction of outer hair cells, produced by noise or by certain antibiotic drugs, but, paradoxically, the ability to discriminate differences in sound intensity is often not affected³.

A key to outer hair cell function may be the presence of the cochlear efferent system. Described originally by Rasmussen in 1946, this neural pathway originates in the superior olivary complex of the brainstem, where incoming sensory information from both ears is combined, and runs back out to the cochlea to terminate

almost exclusively on outer hair cells (see figure). The cochlear synapses are probably cholinergic, but the signal traffic around the loop is certainly too slow to be involved in any sort of neural enhancement of frequency selectivity. Loud sounds can stimulate the system but relatively little is known about what activates it in awake animals.

If, during an acute experiment, the efferent system is stimulated electrically, but at rates that are probably unphysiological, there are alterations in sensory outflow from inner hair cells⁴ or the sound reflected from the cochlea⁵. The consensus interpretation of such results is that the efferent system is modulating the outer hair cells, which in turn are influencing the vibration pattern of the basilar membrane, possibly by an active process. The precise mechanism is not yet clear, but perhaps it is highly relevant that isolated outer hair cells can behave like motile cells^{6,7}: when current is injected into them they

Films of liquid crystal displays

A NOVEL means of generating liquid crystal displays, as used in consumer electronic devices, has been developed at the Liquid Crystal Institute at Kent State University, Ohio. The development is based on a technique for creating bubbles of liquid crystal material *in situ* within thin films of a transparent plastic.

The practical advantages of this form of display are that it will allow the construction of very large displays; and that the need for a polarizing screen in front of the display is obviated. In the published account of the development by J.W. Doane and colleagues, the plastic in which the bubbles are embedded is said to be a resin formed from epichlorohydrin and bisphenol A (*Appl. Phys. Lett.* 48, 269; 1986).

The principle of the new display material is that of the scattering of light from microspherules embedded in an otherwise homogenous transparent material. The trick seems to be to find a nematic liquid crystal material, one in which planar molecules are stacked parallel to each other, whose refractive index in a perpendicular direction is equal to that of the embedding film and whose refractive index in the direction of alignment is greater. One of the materials used for making the liquid crystal inclusions is cyanobiphenyl.

Although there may be some degree of nematic ordering in sufficiently small inclusions, which appear to be about a tenth of a micron in diameter, the random orientation of the birefringent axes of the inclusions will ensure that the light incident at right angles to a plastic film will be diffusely scattered, making the surface appear white. But at places

where there is also an electric field perpendicular to the surface of the film, the degree of nematic order will be increased and, more important, the birefringent axis of the inclusions will be aligned in the same direction. The result is that scattering is reduced, so that the patches concerned will appear black against an apparently white background.

Experiments with a film 24 microns

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Nematic microdroplets, about 0.4 microns in diameter, formed spontaneously by film polymerization (photograph courtesy of J.W. Doane).

thick sandwiched between two glass slides carrying transparent conductors are reported to show that the transition from white to black (or from opaque to transparent) is complete with a potential difference of 30 volts. A possible limitation of the usefulness of the device is that the threshold of the response is not sharp. A potential advantage of this liquid crystal display is that the time needed to convert the system from one state to the other is measured in milliseconds, which compares well with existing liquid crystal displays. This time appears to depend on the shape of the liquid crystal inclusions, an issue to which the authors say they are now devoting themselves. □

appear to elongate or shorten depending on the sign of the current. It remains to be seen whether the efferent system could modify this motility and, more importantly, whether such micromechanical properties can be built into a three-dimensional model of the cochlea to explain its observed macroscopic dynamics.

As part of their interest in cochlear implants, Dodson *et al.* now report² the effects of electrical stimulation, in some cases for a period of up to 20 weeks, by an electrode placed at the base of the guinea pig cochlea. Although they find lesions occurring near the implanted electrode, as reported by other workers³, the surprising observation is that there is degeneration of outer hair cells and efferent terminals in the contralateral cochlea. The authors also show that corresponding basal regions on both sides are affected which, because of the known partial tonotopic organization of the efferent system⁴, demonstrates that the effects do not arise from non-specific trauma.

It thus seems that this type of electrical stimulation can mimic loud sound, known to generate activity in efferent axons. The morphological damage reported by Dodson *et al.* is indirect, although convincing, evidence for abnormal and chronic activation of the efferents. The cochlea is evidently not well-adapted to various prolonged stimuli. But the clinical message is clear: subtle changes may be introduced into the nervous system by long-term electrical stimulation by implants. At the moment, this is a somewhat academic caution, for it seems unlikely that the operation would be attempted on anyone who does not exhibit profound bilateral sensorineural deafness.

The most puzzling feature, as with many reports involving efferent function, is the absence of any obvious physiological correlate of hair-cell destruction. Why should this pattern of chronic activation produce degeneration of the outer hair cells? Does it mean that, normally, the activity in efferent axons is indeed extremely low and that one of the functions of the pathway is to provide some type of trophic regulation of the cells? It seems unlikely that we have heard the final word about the efferent system. Nor for that matter about its target, the outer hair cell. □

1. Parkins, C.W. & Anderson, S.W. (eds) *Ann. N.Y. Acad. Sci.* **405** (1983).
2. Dodson, H.C. *et al. Nature* **320**, 65 (1986).
3. Prosen, C.A. *et al. Science* **212**, 1286 (1981).
4. Brown, M.C. & Nuttall, A.L. *J. Physiol., Lond.* **354**, 625 (1984).
5. Mountain, D.C. *Science* **210**, 71 (1980).
6. Brownell, W.E. *et al. Science* **227**, 194 (1985).
7. Ashmore, J.F. *J. Physiol., Lond.* **364**, 4P (1985).
8. Duckert, L.G. & Miller, J.M. *Acta otolar. Suppl.* **411**, 28 (1984).
9. Warr, W.B. & Guinan, J.J. *Brain Res.* **173**, 152 (1979).

Jonathan Ashmore is in the Department of Physiology, Bristol University Medical School, University Walk, Bristol BS8 1TD, UK.

Ocean Drilling Program

High-latitude palaeoceanography

from the Leg 105 shipboard scientific party*

LEG 105 of the Ocean Drilling Program sought to elucidate the palaeoclimatic and palaeoceanographic changes in the climatically sensitive high-latitude region of Baffin Bay and the Labrador Sea (see figure) during the past 55 million years (Myr). These seas may have constituted an important corridor for exchange of surface- and deep-water masses between the Arctic and Atlantic oceans since as early as the late Cretaceous, and were the only connection between these oceans until the formation of the Norwegian–Greenland Sea during the Eocene.

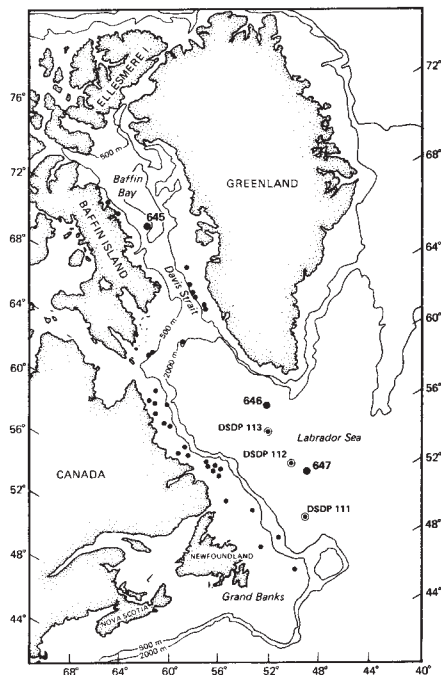
The north–south transect of sites drilled during ODP Leg 105 provides an opportunity for detailed regional palaeoceanographic studies and composes part of an east–west transect that will allow comparisons of the timing and rates of climate and ocean circulation changes over the whole of the North Atlantic and Norwegian Sea region from the Palaeogene to Recent. And the dating of major seismic stratigraphic units will allow more complete understanding of basinwide sedimentation and tectonic events.

Continuous, high-sedimentation rate sequences (ranging from 40 to 130 m per Myr) of Plio–Pleistocene age were penetrated at three drilling sites (645–647 in the figure). These sequences provide the rhythmically sedimented sequences and floral and faunal changes that result from high-frequency climatic oscillations. Leg 105 drilling also confirms that initiation of ice-rafting in the northwestern Atlantic occurred at 2.5 Myr, but indicates that glaciation in the Baffin Bay region may have begun earlier than this.

Climatic deterioration and southwards transport of arctic water masses in Baffin Bay began in the middle Miocene or earlier, and significant cooling of surface waters as early as the late Eocene is indicated by planktonic assemblages from the Labrador Sea. A climatic optimum occurred during the late middle Miocene. Drilling results also demonstrate that building of major sedimentary drifts resulting from intensified deep-water circulation began in the middle to late Miocene in the Labrador Sea.

The age of the ocean crust at site 647 is 55–56 Myr, in agreement with estimates from magnetic anomaly correlations and confirming a tectonic model that requires

a major reorientation of seafloor spreading centres in the Labrador Sea at about that time. Subsidence at Baffin Bay site 645 began at about the same time (63–55 Myr). Extrapolation of sedimentation rates at site 645 to a major deep regional reflector suggests an Eocene–Oligocene age. Because the reflector extends across the entire basin, Baffin Bay must have stopped forming by 36 Myr, the same



Location of sites drilled during ODP Leg 105. Also shown are the locations of DSDP sites and exploratory wells drilled on the continental shelves.

time that spreading ceased in the Labrador Sea.

The first appearance of ice-rafted material at sites 646 (drilled at 3,450 m water depth on the northern flank of Eirik Ridge) and 647 (drilled in 3,860 m of water south of the main part of Gloria Drift) suggests that both events occurred at 2.5 Myr, confirming the earlier results obtained by the Deep Sea Drilling Project (DSDP) Legs 81 and 94 for the northeastern Atlantic.

At site 647 the earliest ice-rafted material rests unconformably above a highly condensed sequence of Upper Miocene strata (5.6–8.2 Myr) which, in turn, unconformably overlies Lower Miocene (17.5 Myr) and older strata. At site 646, the onset of ice-rafted sedimentation closely coincides with the end of a major episode of erosion and depositional modification of Eirik Ridge drift by strong, deep currents. Subsequently the drift was draped

*Co-chief scientists M. Arthur (Univ. Rhode Island) and S. Srivastava (Bedford Inst. of Oceanography); A. Aksu, J. Baldauf, G. Bohrmann, W. Busch, T. Cederberg, B. Clement, M. Cremer, K. Dadey, A. De Vernal, J. Firth, F. Hall, M. Head, R. Hiscott, R. Jarrard, M. Kaminsky, D. Lazarus, A.-L. Monjanel, O. Nielsen, R. Stein, F. Thiebault, J. Zachos and H. Zimmerman.

and smoothed by sediment deposited in part by deep currents. A picture emerges of strong bottom currents sweeping the sea floor during the early to late Pliocene, with the combination of increased supply of terrigenous sediment resulting from glacial activity and weakened deep currents leading to a major change in style of sedimentation at about 2.5 Myr.

The appearance of siliceous planktonic faunas at about 3.5 Myr in site 646 signal the change to colder surface-water masses as a prelude to the glacial epoch. Sediments from site 645 (2,020 m water depth on the continental slope of Baffin Island) record the onset of glacial ice rafting at this time. An age of 2.5 Myr can be assigned to the major flood of ice-rafting. Isolated pebbles, mostly shale, are fairly common in sediments as old as 3.5 Myr, and a few such clasts were observed in strata as old as 8.0 Myr. Although not definitive, the evidence suggests an earlier inception of glacial activity in the Baffin Bay region than in the northern Atlantic. The 340 m of glacial strata younger than 2.5 Myr at site 645 overlie an important regional unconformity and thicken considerably towards, and onlap the unconformity from, the west Greenland margin. Seismic evidence suggests that a major influx of clastic material began significantly earlier than 2.5 Myr in association with initial build up of the Greenland ice sheets. The Leg 105 drilling results and seismic data demonstrate that similar floods of terrigenous sediments occurred in the Labrador Sea several million years before the generally accepted time of Northern Hemisphere glaciation.

The Leg 105 results also provide critical information on the older history of bottom-water circulation in the region. Although there is seismic and oceanographic sup-

port for substantial influence of deep currents on sedimentation at the drill sites, there are few sedimentary structures or textural features in the cores that indicate current winnowing or redeposition — much of the sediment is bioturbated homogeneous. At sites 646 and 647, a regional reflector, termed R4 and tentatively correlated with a similar reflector of Eocene/Oligocene or early Oligocene age in the northeastern Atlantic, was penetrated. The Atlantic reflector was thought to represent an episode of bottom-water formation and the resulting seafloor erosion. The reflector at site 646 occurs well below the base of drift deposits recognizable in seismic records, but unexpectedly is of late Miocene age (about 10 Myr). Therefore, major drift sedimentation on Eirik Ridge must have occurred later, beginning after about 6 Myr. R4 is early Oligocene in age (34 Myr) at site 647. Here it represents a change from dominantly calcareous pelagic sediments below to siliceous-biogenic pelagites above, but has no obvious association with erosion or sedimentation caused by bottom currents. However, a major change in the benthic foraminiferal fauna occurred in the late Eocene to earliest Oligocene, indicative of a change in bottom-water mass characteristics.

None of the evidence excludes intensification of deep circulation in the latest Eocene–Oligocene, but the drilling results and seismic data show that the most significant change in intensity occurred after mid-Miocene. An unexpected result from the drilling in Baffin Bay is the mid-Miocene increase in southerly transport of arctic water masses and intensification of deep circulation; seismic reflection profiles show that drift deposits of that age occur on the Baffin Island slope. □

(Oxford Scientific Films)

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Anna's hummingbird (*Calypte anna*)

rate for sluggish *Homo sapiens*⁴, and similar to the 57 Kcal kg⁻¹ h⁻¹ for a shrew (*Sorex spp.*) of similar size⁵. Small homeotherms, and birds in particular, have very high mass-specific metabolic rates and therefore have to spend much of their day feeding just to stay alive. In a classic paper, John Gibb recorded that a 9.0-g insectivorous bird such as the coal tit (*Parus ater*) has to eat one insect the size of an aphid every 2.5 s (or one the size of a small lepidopteran larva every 25 s) throughout daylight hours to keep alive⁶.

The Anna's hummingbird fuels its prodigious energy needs with sugar-rich nectar: it eats about 180 meals per day, totalling roughly three times its own body-weight of fluid. As might be expected, hummingbirds are very efficient at collecting nectar. Behavioural studies of the sequence in which they visit flowers on a plant⁷, the number of flowers they defend⁸ and the size of the meals they eat⁹, have shown consistently that these birds adopt strategies of feeding that come close to maximizing net rate of energy gain.

Diamond *et al.* show that hummingbirds also have a remarkably efficient digestive system for extraction of sugars from nectar. First, the throughput of food is very fast: markers appear in the faeces only 15 min after ingestion, and the average retention time is only 49 min. Second, despite this rapid throughput, hummingbirds can extract about 97 per cent of the glucose ingested in a meal. This high rate of absorption is achieved by an exceptionally high maximum velocity of active transport together with a very low passive permeability of the gut wall, which is perhaps necessary to prevent loss of blood solutes to the fluid flowing through the gut.

Avian physiology

Busy doing nothing — efficiently

from John R. Krebs and Paul H. Harvey

THE statement that animals in the wild spend much of their time doing nothing is technically incorrect — as Niko Tinbergen used to point out to his students, only a dead animal does nothing. But it is true to say that animals of many species spend a considerable amount of time, even during their so-called active phase of the 24-hour cycle, sitting or otherwise inactive. And a report on page 62 of this issue¹ shows that this is true even in unexpected circumstances.

It is not surprising that large herbivorous animals such as sheep or cows spend much time inactive. For contented cows, the limiting factor in meeting daily food requirements is the time taken to break down the relatively indigestible plant mat-

ter that makes up the diet. Ingesting food can be quite rapid, but processing it in the gut is a long, slow business. The same is true for wood pigeons living on a low-quality diet of brassica, although not when they feed on more nutritious clover².

Much more surprising is the new observation by Diamond *et al.*¹ that hummingbirds in the wild spend up to 75 per cent of the day sitting on a perch. Hummingbirds have one of the highest mass-specific metabolic rates in the animal kingdom because of their small body mass and energetically costly method of feeding by hovering in front of flowers to suck nectar. A 4.8-g Anna's hummingbird (*Calypte anna*) burns energy at a phenomenal 58 Kcal kg⁻¹ h⁻¹ (ref. 3), about 30 times the

Given their high energy needs and great efficiency of acquisition and extraction of energy, why do hummingbirds spend so much time sitting? Diamond *et al.* suggest that the rate-limiting step in digestion for the birds is the time taken for the crop (the specially modified region of the oesophagus used to store food immediately after ingestion) to empty. They used a dilution technique with radioactively labelled polyethylene glycol to estimate crop contents at the start of a meal and after known intervals, and show that it takes about 4 min for the crop to half-empty after a 100- μ l meal. During this 4-min period, while the bird is waiting to make space for its next meal, it sits on a perch and minimizes its energy expenditure. The observed pauses between meals are usually about 4 min, as predicted by the hypothesis of Diamond *et al.*

Are digestive bottlenecks, like the one so elegantly demonstrated by Diamond *et al.* for hummingbirds, found in other small vertebrates? It is well known that shrews

spend much of their time feeding, but it is less often mentioned that during an activity period, and shortly after feeding, they frequently rest for 5 min at a time (I. Hanski, personal communication). Perhaps for these animals, like hummingbirds, rates of digestion limit rates of ingestion⁵. Optimization processes evolve under certain constraints, and rates of digestion may provide a more widespread constraint to activity patterns than previously thought. $\square$

1. Diamond, J.M. *et al.* *Nature* **320**, 62 (1986).
2. Kenward, R.E. & Sibley, R.M. *J. appl. Ecol.* **14**, 815 (1977).
3. King, J.R. in *Avian Energetics* (ed. Paynter, R.A.) 4 (Nuttall Ornithological Club, Massachusetts, 1974).
4. Ferro-Luzzi, A. in *Energy and Effort* (ed. Harrison, G.A.) 115 (Taylor & Francis, London, 1982).
5. Hanski, I. *Ann. Zool. Fennici* **21**, 157 (1984).
6. Gibb, J. *Ibis* **102**, 193 (1960).
7. Pyke, G. *Am. Zool.* **18**, 739 (1978).
8. Carpenter, F.L., Paton, D. & Hixon, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7259 (1983).
9. DeBenedictis, P.A., Gill, F.B., Hainsworth, F.R. & Wolf, L.L. *Am. Nat.* **112**, 301 (1978).

John Krebs and Paul Harvey are in the Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

Oceanography

Conductivity of the sea floor

from F.E.M. Lilley

THE effort in earth science to move from observations on continents to observations on seas and sea floors has produced spectacular results; indeed, modern geology is based largely on the rewards from such adventures. The deep ocean is a remote and hostile place in which to work, requiring the development of special equipment and techniques. Once these difficulties are overcome, the deep sea floor has several advantages. Large areas of it are flat and uncluttered; it is a generally undisturbed and stable environment; equipment can be left without interference; and there seems to be little difficulty about getting owner permission — indeed there is great attraction in the internationality of seafloor science. The recent work by Cox and co-workers reported elsewhere in this issue (*Nature* **320**, 52; 1986) represents a significant step in

efforts to sound the sea floor electromagnetically.

Data on seafloor seismic velocities, magnetic patterns, sediment thicknesses and heat flow, which have proved so fertile for earth science, are valuably embellished by information on electrical conductivity. At the seafloor interface this parameter relates to the porosity and thickness of the sediments; below the sediments it reflects water content in the oceanic lithosphere with implications for fracture patterns and the possibility of seawater convection, as well as the possible inclusion in the lithosphere of mineralogy of high electrical conductivity. Still deeper, at asthenospheric depths, the condition of the asthenosphere itself — is it partially molten? — should be evident from an electrical conductivity profile.

In addition to its obvious geological relevance, electromagnetic measurements promise to yield much oceanographic information, and for a correct interpretation of such measurements a good understanding is necessary of the leakage electric currents that flow through the sea floor. These currents depend in turn on the seafloor electrical conductivity.

The work by Cox *et al.* reported in this issue describes a controlled-source electromagnetic experiment on the Pacific sea floor off California. The method (see figure) is an advance on the previous experiment of P.D. Young and C.S. Cox (*Geophys. Res. Lett.* **8**, 1043; 1981), in particular because of a new receiver arrangement (see Webb S.C., Constable S.C., Cox C.S. & Deaton T.K. *J. Geomag. Geoelect.* **37**, 1115; 1985). The new equipment allows a greater separation of transmitter and receiver, giving successful transmission over a distance of 65 km.

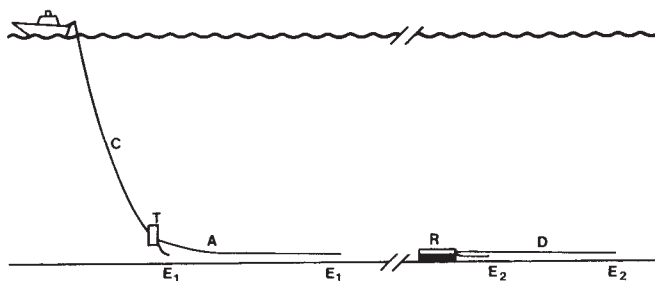
Several receivers are put out on the sea floor, and the ship steams away from them, towing the transmitting cable. The transmitted signal is rapidly attenuated in seawater, but is detected at the receivers having followed paths through the more weakly conducting sea floor (see sketch). Later the receivers are retrieved, having floated to the surface.

These pioneering results show that below a sediment layer, the crustal layer, of thickness some 5 km, has a conductivity of 10^{-3} S m^{-1} , and is underlain by a thicker region of very low conductivity — less than $2 \times 10^{-5} \text{ S m}^{-1}$. In rock material, such low conductivity means dry conditions, and Cox *et al.* interpret these particular results to mean a low water content (less than 0.1 per cent by volume) in the upper mantle.

The results of further seafloor recordings will be awaited with interest not only from the Scripps team, but also from other groups with electrical techniques now under development — for example the Canadian magnetometric offshore electrical sounding (MOSES) procedure, whose name is appropriate because the technique effectively 'parts the waters' to reveal the sea floor beneath. The international EMSLAB experiment had a successful observing period last (northern) summer. EMSLAB is an electromagnetic study of the lithosphere and asthenosphere beneath the Juan de Fuca Plate and the adjacent continent, with a superarray of instruments recording natural source fields, both onshore and offshore in western North America and the Pacific Ocean. The combination of natural-source results, such as those of EMSLAB, with the controlled-source results that I have described here, will provide mutually beneficial data.

Finally, the experiment of Cox and colleagues demonstrates the reception of a weak electromagnetic signal across 65 km of deep sea floor by transmission through a resistive rock path under the ocean water. The idea of communication in this way between two sites on the ocean floor has aspects yet to be fully explored. $\square$

F.E.M. Lilley is a fellow in the Research School of Earth Sciences, Australian National University, PO Box 4, Canberra, ACT 2601, Australia.



Transmitter and receiver on the sea bed. C, Cable bringing power from ship; T, transformer housing; A, insulated wire transmitter antenna (600 m); E₁, transmitter electrode (bared wire); R, receiver housing; D, receiver antenna (600 m); E₂, receiver electrodes (Ag-AgCl).

Inflammation

The mediators of steroid action

from R. J. Flower

In the clinic, steroids are widely used for the treatment of many types of inflammatory disease. Although steroids interrupt the inflammatory response at many levels, it is not known precisely how they accomplish this and they can have serious side effects. On page 77 of this issue¹, Barbara Wallner and her colleagues at Biogen report an important step in dissecting the mechanisms of steroid action — they have succeeded in identifying the human complementary DNA for lipocortin, an endogenous anti-inflammatory substance, whose production is regulated by circulating steroids and is intimately linked to their mechanism of action.

The lipocortins² are a family of related proteins, independently discovered by several groups³⁻⁸, with the unusual property of blocking the enzyme phospholipase A₂. This enzyme is important for several reasons. First, it is involved in the signal transduction process, which enables leukocytes to respond to noxious stimuli, causing them to migrate in chemotactic gradients and to effect a repertoire of responses crucial to the development of inflammation and other host defence mechanisms.

Second, by catabolizing membrane phosphatides, phospholipase A₂ liberates much of the substrate used for the generation of inflammatory mediators such as the prostaglandins, leukotrienes and platelet activating factor. Preparations containing lipocortin can suppress the production of mediators by cells, prevent neutrophil and macrophage chemotaxis and suppress inflammation itself in animal models. Although the proteins of the lipocortin family have been fairly well characterized biologically, and even highly purified in small amounts, Wallner *et al.*¹ present the first structural data.

The authors first purified lipocortin from rat peritoneal lavage fluid (a rich source of this protein) and used the amino-acid sequence information from

the tryptic fragments obtained to synthesize oligonucleotide probes which were found to hybridize with messenger RNA extracts from both rat and human cells. One of these probes recovers from a human lymphoma-derived cDNA library a DNA sequence that encodes a protein containing the sequence of all the tryptic fragments of rat lipocortin. This gene was expressed in *Escherichia coli* and the protein thus obtained was found to be a monomer of relative molecular mass (M_r) about 37,000, corresponding closely to a commonly observed species of M_r 40,000 (the discrepancy could be caused by the absence of glycosylation: a potential glycosylation site has been identified).

Lipocortin has several notable features. The first is the highly polar nature of the molecule: almost a third of its amino acids are charged and these seem to be distributed throughout the length of the molecule. This fact is perhaps surprising. The expectation of most workers was that the protein would be mainly hydrophobic as might befit a membrane-associated protein involved in the control of lipid metabolism. Another feature is the high degree of homology and immunological cross-reactivity which exists between the human and the rat protein, suggesting that this molecule is highly conserved.

Wallner *et al.*¹ have examined the occurrence of lipocortin mRNA in the rat and find it widely distributed but most abundant in macrophages, lung, thymus and spleen. These findings correlate quite well with the previously observed distribution of anti-phospholipase activity.

As one might expect, treatment of rats with dexamethasone, a synthetic steroid, led to a rapid increase in lipocortin mRNA in peritoneal macrophages (treatment of these cells *in vitro* works just as well), although there seems to be a slight anomaly in that in at least one set of experiments this was not followed by the expected rise in cell lipocortin levels.

The inhibitory activity of lipocortin on pancreatic phospholipase A₂ was tested *in vitro* on a substrate of autoclaved *E. coli* prelabelled with fatty acids. The recombinant material is extremely active in this system, and comparable to the naturally occurring protein.

The blood of all healthy mammals contains circulating glucocorticoid steroids and the toe of the dose-response curve of steroid-stimulated lipocortin release lies within this physiological range. It is therefore not surprising that lipocortin is present in normal cells and tissues. There has been a growing realization among workers

in the lipocortin field that this protein is not simply involved in the 'crisis management' of inflammation, but almost certainly regulates the activity of membrane-bound phospholipases under resting conditions as well.

To do this effectively, there needs to be a mechanism whereby phospholipase can be relieved of lipocortin inhibition at times when its activation is essential to bring about a cellular response. One way in which this might occur has been suggested by Hirata's work on lipocortin phosphorylation⁶. He has demonstrated that this protein can be phosphorylated by protein kinases and that this results in a loss of its inhibitory activity which can be subsequently reversed by treatment of the protein with phosphatases. As so many stimuli seem to exert their actions by interacting with cell-surface receptors and causing activation of protein kinases, this mechanism would seem to be of great topical interest. Wallner *et al.*¹ find that lipocortin has two potential phosphorylation sites (on tyrosine and threonine residues) which would fit nicely with Hirata's notion, although it is not yet clear whether phosphorylation leads to inactivation of the recombinant molecule.

The significance of the paper from Wallner *et al.* lies not so much with the sequencing and cloning of lipocortin as such but in the many avenues now open to us. We obviously want to know the minimum effective inhibitory sequence of lipocortin and as much as possible about the way in which the protein inhibits phospholipase so that drugs can be designed that have the same effect.

The wide spectrum of side effects preclude the use of steroids for long periods of time except in the very seriously ill. These side effects are caused by changes in the transcription of specific genes in the same way as the anti-inflammatory effects. It has long been an article of faith of scientists working in this area that if we could identify and isolate the 'second messengers' of steroid action which are responsible for the anti-inflammatory effects, we should also be able to produce drugs which possess many of the beneficial actions of steroids without incurring the heavy penalty of side effects. The real value of the work by the Biogen team is that it enables us to make headway in that direction. □

1. Wallner, B.P. *et al.* *Nature* **320**, 77 (1986).
2. Di Rosa, M. *et al.* *Prostaglandins* **28**, 441 (1984).
3. Flower, R.J. & Blackwell, G.J. *Nature* **278**, 456 (1979).
4. Blackwell, G.J. *et al.* *Nature* **287**, 147 (1980).
5. Hirata, F. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 2533 (1980).
6. Hirata, F. *J. biol. Chem.* **256**, 7730 (1981).
7. Russo-Marie, F. & Duval, D. *Biochim. biophys. Acta* **712**, 177 (1982).
8. Cloix, J.F. *et al.* *Br. J. Pharmacol.* **79**, 313 (1983).

R.J. Flower is Professor of Pharmacology in the School of Pharmacy and Pharmacology, University of Bath, Claverton Down, Bath BA2 7AY, UK.

100 years ago

M. PASTEUR, at the last sitting of the Paris Academy of Sciences, stated that out of 325 cases of inoculation for hydrophobia, only one had failed — namely, that of the youth Peltier, who came too long after being bitten, and under very unfavourable conditions. He advocated the establishment of an international hospital, to which patients would come from all parts of the world; and he suggested a discussion as to the locality and the fund for its support. Prof. Pasteur announced that he should next investigate whether diphtheria could not be treated by a similar process.

From *Nature* **33** 423, 4 March 1886.

Probes in Huntington's chorea

SIR—Over two years ago *Nature* published the first linkage of a marker with Huntington's chorea and underlined its importance and significance with a full and lucid annotation.

Under the expectation that the probe presented would be of some clinical use under certain conditions and that it would be followed by closer probes, some of which would bracket this locus, we organized a register of cases in this Region (population 2,500,000) and acquired funds and limited facilities allowing us to start to apply DNA techniques to the practical problems presented by patients and their relatives. We now know of over 300 cases in 140 families, and laboratory work started in October. We are not inexperienced and many members of the families we are trying to help are responsible and intelligent.

On asking for the probe we received a letter from Dr Gusella stating that there were ethical problems in its use, that further research on heterogeneity was needed and that we could only have it if we did not use it clinically. We are of course aware that on its own, it is too far from the locus to be of clinical value in most families. We are also aware that cases are often misdiagnosed; and that this is sometimes euphemistically included under heterogeneity.

No novel ethical problems are involved in initiating work by the cautious application of new and unreliable techniques. Any relevant test on a patient when the diagnosis is in doubt must be followed by news which is either good or bad. Suicide in response to bad news is a rational anxiety but rare in medical practice. If Wasserman had published his test for syphilis, which was far from reliable, in a way which delayed its application, neurosyphilis might now be more common than Huntington's chorea. This infectious disease involved far more difficult problems in handling patients and their families; many individuals must have been distressed by investigations based on error and there were probably a few suicides. But the disease is now rare in Northern Europe. In no field of effective medicine can techniques be applied without casualties.

Quite apart from these serious ethical issues now arising in the field of DNA, *Nature* is a scientific journal and it should be assumed that others will be able to confirm the claims made in articles it publishes. The time involved in preparing a paper of sufficient standard is so long that there is little justification for withholding information or resources. It would hardly be considered ethical to publish an article on an interesting star, but to withhold its

position to allow a monopoly of further research or development. In this new field *Cell* has already given a lead — perhaps *Nature* could follow and neither publish nor refer to work which cannot be validated.

D.C. WATT
R.H. LINDENBAUM
J.A. JONASSON
J.H. EDWARDS

*Medical Genetics Unit,
Churchill Hospital,
Headington,
Oxford OX3 7LE, UK*

GUSELLA REPLIES—Watt *et al.* appear to confuse scientific research with clinical practice. Our report in *Nature* did not present a presymptomatic diagnostic test for Huntington's disease (HD), but demonstrated linkage of HD to a genetic marker defined by the DNA probe G8 on chromosome 4¹. We agree wholeheartedly that if scientific reports are published it should be possible to repeat and confirm them. To this end, we have distributed the G8 probe to over 30 laboratories in 10 countries. We are also collaborating with many other investigators to apply the probe to additional HD families. In all cases, with the agreement of the investigators involved, the work is being performed for research purposes only, not to derive predictive information for delivery to "at risk" individuals.

Our reasons for limiting clinical use of the marker were referred to in the original report of its discovery in *Nature*. There is considerable precedent in genetics, including human genetic disease, for non-allelic heterogeneity in which mutations at two or more separate loci result in the same phenotype. If a second unlinked locus were the cause of HD in some families, predictive testing for these using the G8 marker would be no better than guesswork, and many "at risk" family members might become unnecessary "casualties" to false information.

Heterogeneity can only be ruled out by typing individuals of defined HD status in fairly large pedigrees, unlike those described to us by Dr Edwards and his colleagues. When we received the request from this group to proceed with predictive testing, linkage of the marker to the disease locus had only been published for four independent families, far too few to conclude that a second HD locus does not exist in a significant proportion of families^{1,2}. Efforts to test the possibility of heterogeneity are ongoing in many centres and should provide conclusive results in the near future. We have not performed any presymptomatic or prenatal diagnoses using the marker and feel that universal clinical application of the G8 probe in HD remains premature.

Unlike Watt *et al.*, we feel that there are

significant considerations in the presymptomatic diagnosis of HD that make it similar to several other neurological and behavioural disorders for which markers may soon be found, but that distinguish it from a communicable disease like syphilis. HD is a late onset condition that affects the victim both physically and mentally, but can only be transmitted to progeny. In addition to motor disturbance, psychiatric symptoms are common, and the risk of suicide is significantly increased^{3,4}. As there is no effective therapy for the disorder, any presymptomatic testing is purely informational and will not affect treatment. We believe that when accurate presymptomatic diagnosis is possible, it should be made available for the benefit of those "at risk" individuals who request it, rather than for the purposes of physicians, other family members, insurance companies, employers, or governments. A positive test would indicate to an apparently healthy individual that he or she has a very high probability of becoming progressively disabled, both mentally and physically, later in adult life, and that any children might await a similar fate. In doing what is best for the patient, it is the obligation of the healthcare professional to deliver this disruptive and potentially devastating information in a manner which minimizes its many possible detrimental effects on both the individual and his or her family.

When accurate presymptomatic testing is possible, the appropriate setting will be one which includes adequate provisions for pre- and post-test counselling as well as long-term psychological support⁵. We have been working closely with the HD voluntary organizations who represent patients and their families to ensure that the needs of the "at risk" individuals will be met when presymptomatic testing can be initiated. We have formally raised these issues for consideration at international meetings of the World Federation of Neurology Research Group on Huntington's Disease and of the International Huntington's Association and have also privately solicited input from many expert neurologists, psychiatrists and geneticists.

It is our intention that when accurate presymptomatic testing is possible, the G8 probe will be made available to any qualified centre through the Massachusetts General Hospital Committee for Presymptomatic HD Testing. In the meantime, we regret having to place any restrictions on the use of the G8 probe, but these are not ones which should delay scientific progress or test development. We and others are actively searching for additional markers that will increase both the accuracy and applicability of the potential presymptomatic test, but the heterogeneity issue must be resolved before proceeding to preclinical diagnosis. We feel a mor-

al responsibility to prevent the premature clinical use of the G8 probe until the information derived from it can be judged accurately. A scientist cannot ignore the social consequences of his work, especially in medicine, and should not abrogate his social responsibility.

JAMES F. GUSELLA

Neurogenetics Laboratory,
Massachusetts General Hospital,
Boston,
Massachusetts 02114, USA

1. Gusella, J.F. *et al.* *Nature* **306**, 234-238 (1983).
2. Folstein, S.E. *et al.* *Science* **229**, 776-779 (1985).
3. Schoenfeld, M. *et al.* *J. Virol. Neurosurg. Psychiat.* **47**, 1283-1287 (1984).
4. Farrer, L. *Am. J. med. Genet.* (in the press).
5. Wexler, N.S. *Archs Neurol.* **42**, 20-24 (1985).

Industrial age leading to the greening of the Earth?

SIR—That the concentration of CO₂ in the atmosphere has increased from 265 parts per million (p.p.m.) in pre-industrial times to 314 p.p.m. in 1958 and more than 345 p.p.m. today has been established by several means, but little attention has been paid to the hypothesis that the enrichment of atmospheric CO₂ may greatly increase plant growth and development^{1,2}.

A good test of this hypothesis^{3,4} is to be found in the secular variation of the seasonal CO₂ cycle amplitude. Does it increase with time in response to an increasing plant growth and decay cycle caused by the rising atmospheric CO₂ concentration?

Recent years have seen several independent confirmations of this scenario (Fig. 1). The three analyses for Mauna Loa, Hawaii, yield very similar slopes and the two analyses for Weather Station P yield essentially identical slopes, evidence which points towards the reality of the phenomenon.

Further support for this view is provided by the latitudinal plot of the mean result for each site shown in Fig. 2. The

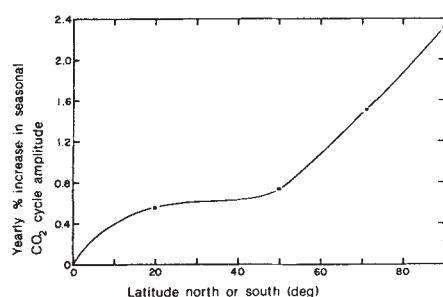


Fig. 2 The mean yearly percentage increases in seasonal CO₂ cycle amplitudes of the four sites represented in Fig. 1 plotted against their latitudinal locations irrespective of north/south designations.

line connecting the data points of this figure was drawn by eye and forced to intersect the origin. A negligible change in the seasonal CO₂ cycle amplitude in this region is predictable, since the area in the immediate vicinity of the Equator is a mixing zone for the out-of-phase effects of the diametrically-opposed seasonalities of the Northern and Southern Hemispheres. Since this mixing and counterbalancing of effects should decrease both north and south from the Equator, however, we would expect the yearly percentage increase in the seasonal CO₂ cycle amplitude to monotonically rise with increasing latitude, as demonstrated in Fig. 2.

In light of this body of evidence, it is difficult to reject the conclusion that the predicted CO₂-induced "greening of the Earth" is upon us. Pearman and Hyson⁵, for example, concluded that "it is most probable that there has been an increase in the summer net ecosystem production of the Northern Hemisphere". Similarly, Cleveland *et al.*⁶ concluded that "it is most likely that the CO₂ seasonal behaviour reflects an increase in global photosynthetic activity". Finally, both Bacastow *et al.*⁷ and Keeling *et al.*⁸ have concluded, respectively, that it seems likely that the

increasing seasonal CO₂ cycle amplitude "mainly reflects enhanced metabolic activity of the land biota" and "suggests a heightening plant activity".

May I thus suggest that the elusive "first detection" of global CO₂ effects has in fact been accomplished, and that, as predicted by Wittwer³, it is "biological, rather than climatic".

S.B. IDSO

US Water Conservation Laboratory,
4331 E. Broadway,
Phoenix, Arizona 85040, USA

1. Kimball, B. A. *Agron. J.* **75**, 779 (1983); *Carbon Dioxide and Agricultural Yield: An Assemblage and Analysis of 770 Prior Observations* (US Water Conservation Laboratory, Phoenix, Arizona, 1983).
2. Lemon, E.R. (ed.) *CO₂ and Plants: The Response of Plants to Rising Levels of Atmospheric Carbon Dioxide* (Westview, Boulder, Colorado, 1983).
3. Wittwer, S.H. *New Scientist* **95**, 233 (1982).
4. Idso, S.B. *Carbon Dioxide: Friend or Foe?* (IBR, Tempe, Arizona, 1982).
5. Pearman, G.I. & Hyson, P.J. *geophys. Res.* **86**, 9839 (1981).
6. Cleveland, W.S., Freeny, A.E. & Graedel, T.E. *J. geophys. Res.* **88**, 10934 (1983).
7. Bacastow, R.G., Keeling, C.D. & Whorf, T.P. *J. geophys. Res.* **90**, 10529 (1985).
8. Keeling, C.D., Whorf, T.P., Wong, C.S. & Bellagay, R.D. *J. geophys. Res.* **90**, 10511 (1985).

How slow is protein conformation change?

SIR—I have recently performed a calculation that leads to a surprising conclusion. The rate at which pepsinogen exposes its binding site during its acid-initiated activation has been measured (Glick *et al.*, *Biochemistry*, in the press). It is a first-order event with a $\tau_{1/2}$ of about 1 s. If one assumes that the conformational change involved in this transformation requires a displacement of 10 Å of the amino-terminal segment (which subsequently will be cleaved off), the rate of movement is calculated to be 2×10^{-9} km per h. This rate seems incredibly slow; but to what may one compare it?

Using the relation $\bar{x} = (2Dt)^{1/2}$, one can calculate a diffusion coefficient; it is 1.25×10^{-15} cm² s⁻¹, an exceedingly low value when seen in comparison to the range of diffusion coefficients measured for proteins (10^{-6} to 10^{-7} cm² s⁻¹), not to say for peptides of the size of the activation sequence of pepsinogen. It would seem, therefore, that the movement of the activation sequence is not a smooth steering of the amino-terminal segment into its new configuration. During activation, the amino-terminal sequence blindly gropes here, randomly breaking a salt bridge, forming a hydrogen bond there, until the system nestles in an energy minimum along the reaction pathway.

It would be interesting to know if any of your readers can make similar calculations in systems with which they are familiar.

DAVID M. GLICK

Department of Biochemistry,
Medical College of Wisconsin,
Milwaukee, Wisconsin 53226, USA

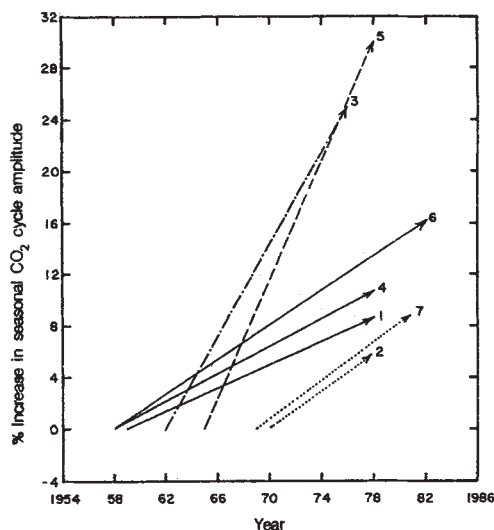


Fig. 1 Linear percentage increases in the seasonal atmospheric CO₂ cycle amplitudes observed at several locations over the time spans indicated by the numbered arrows. (1) Mauna Loa, Hawaii⁵, (2) Weather Station P in the North Pacific⁵, (3) Point Barrow, Alaska⁵, (4) Mauna Loa, Hawaii⁶, (5) South Point Barrow, Alaska⁵, (6) Mauna Loa, Hawaii⁷, (7) Weather Station P⁸.

Darwinism and dialectics

David L. Hull

The Dialectical Biologist. By Richard Levins and Richard Lewontin.
Harvard University Press: 1985. Pp.303. \$20, £16.95.

RICHARD Levins and Richard Lewontin are two of the most knowledgeable and innovative evolutionary biologists working today. They also view themselves as Marxist revolutionaries. As Marxists, Levins and Lewontin insist that the economic substructure of a society strongly influences its ideational superstructure, including science. Because they themselves were educated in a capitalist country, a mechanistic, reductionistic, positivistic ideology quite naturally dominated their own academic education and continues to pervade their intellectual environment. Even so, during 25 years of collaboration, Levins and Lewontin have striven to overcome the influences of their youth and to replace them consciously with a commitment to materialistic dialectics. Because the biology that pervades their book is primarily evolutionary biology, perhaps a more accurate title might have been *The Dialectical Evolutionary Biologist*.

Levins and Lewontin warn that every "theory of the world that is at all powerful and covers a large domain of phenomena carries immanent within itself its own caricature" (p.65). The caricature immanent in Marxism is economic determinism — the attempt to "explain the smallest detail of human history as a direct consequence of economic forces" (p.65). Intellectuals are well aware that all sorts of half-articulated preferences and assumptions influence what they believe. For example, some of the resistance to Darwin's theory in the nineteenth century resulted from how messy and wasteful it was; whatever the God of the Galapagos might be, he certainly was not the Protestant God of Waste Not, Want Not. Intellectuals begin to become restive only when the influence of such factors as reason, argument and evidence is totally discounted, as if our beliefs are dictated solely by social forces. Levins and Lewontin hold no such extreme views. After all, they were able to resist the dictates of their own society; if they can do it, so can others.

Darwinism, too, is a powerful theory of the world. Just as economic determinism is the caricature immanent in Marxism, the "element that is both central to the evolutionary world view and yet so powerful that it can destroy Darwinism as a testable theory is adaptationism" (p.66). Levins and Lewontin are the first to admit that many — probably most — of the features that characterize organisms arose through selection. They object only to the

facility with which adaptations can be attributed to particular traits and the difficulty of testing such hypotheses. Such mistakes would be of only parochial interest to evolutionary biologists if it were not for the extension of adaptationist explanations to the human species. According to

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

"Isms" of explanation — Darwin (top) and Marx.

certain versions of evolutionary theory currently popular, human beings are unlikely to be able to form the genuinely egalitarian, altruistic and cooperative societies that Marxists hope to construct.

As dialectical biologists, Levins and Lewontin share with Darwinians a commitment to materialism and the universality of change but reject the reductionism which they see as pervading present-day Darwinism. As materialists, Levins and Lewontin oppose reference in science, not to abstractions as such, but to reified abstractions. The difference here is that idealists see natural phenomena as the result of real relations among ideal types, while materialists insist on actual material relations occurring among actual material

objects. One such distinction is between archetypes and ancestors. Idealists abstract archetypes from the character distributions that they observe among organisms, while materialists insist that these archetypes play no role in natural processes. Ancestors are organisms, not archetypes. As quaint as reference to archetypes might seem, this cast of mind is not extinct today among biologists; it was, for instance, one source of the protracted debate over cladistics that occurred in the pages of this journal a few years ago.

Although Levins and Lewontin are committed to the "universality" of change, they admit that there are physical constants that do not change: "Of course the speed of light is the same under socialism and capitalism, and the apple that was said to have fallen on the Master of the Mint in 1664 would have struck his Labor Party successor three-hundred years later with equal force" (p.4). They merely oppose the tendency to see too many aspects of nature as being eternal and immutable. The processes governing biological evolution may be changeless, but species are not. Idealists view species as being composed of organisms that are all essentially the same — except of course for a few abnormal, deviant individuals. Levins and Lewontin protest that heterogeneity, not homogeneity, is of the essence of species; more often than not greater genetic heterogeneity exists within a species than between closely related species. Evolution is the process by which rare genes become commonplace and abnormalities are converted temporarily into normalities. Idealists view variation as only an epiphenomenal gloss overlying essential similarity, while for evolutionary biologists variation is the norm and immutability is an illusion. The implications of this general view about species for one particular species, *Homo sapiens*, are obvious.

Because materialism is often associated with reductionism, Levins and Lewontin take great pains to disassociate themselves from the evils of the latter. They view the world as being hierarchically organized. The entities at any one level not only are interrelated but also interpenetrate each other. Organisms are not atomistic entities that interrelate with each other and an autonomous environment; rather they interact with their environment and change it. Nor are the entities at any one level solely a function of the interrelations among entities at lower levels. Entities at all levels possess emergent properties that are not simply a function of lower level phenomena — "there is no basement" (p.278).

It is always difficult to know how to react to the sorts of general philosophical positions urged by Levins and Lewontin. Reductionism does seem overly simplistic

Mansell Collection

and crude, but time and again progress has been made by scientists adopting a reductionist strategy of the most extreme form. Only when Mendelian geneticists divorced themselves from embryology were they able to make any headway in understanding hereditary phenomena. Only now, 80 years later, are biologists beginning to understand how genomes produce phenomes, and this progress would have been impossible without the advances made by Mendelian geneticists. As simplistic as treating evolution as nothing but changes in allelic frequencies may be, this was the simplifying assumption that opened up population genetics. Not all intellectual Philistines make seminal contributions to science, but at times it seems as if only Philistines ever succeed. When scientists get too sophisticated, they become immobilized.

One of the main messages of dialectical materialism is that the world *can* be changed. Like Marx, Levins and Lewontin are not content merely to interpret the world. Several of the chapters in their book deal with agricultural practices and policies, especially in underdeveloped nations. Levins and Lewontin decry the turning of the products of scientists into commodities. Agricultural policies in capitalistic countries are directed primarily at making a profit and only incidentally with feeding the hungry, while socialists tend to institute agricultural policies before the farmers who must implement them have been sufficiently "revolutionized". The wholesale collectivization of Russian agriculture after the revolution is one case in point. The Lysenko debacle is another. The situation seems hopeless — farmers in the Third World are caught between those who simply import agricultural practices that have proved successful in the West, even when the local conditions are totally un conducive, and Maoist ideologues who insist that Lysenko was a "great upholder of materialist method of investigation" (p.189). I sympathize with those who would like to alleviate the suffering in the world, but I find it difficult to see how the downtrodden can survive much more help.

Dialectical materialists object to viewing the world in terms of ideal types, yet they themselves exhibit a strong tendency to treat both Marxism and capitalism as reified abstractions. Perhaps an ideal Marxist state would be Heaven on Earth, but the ideal Marxist state is no more likely to be realized here on Earth than is the ideal capitalist state. Like it or not, we live in the real, not ideal world. Regardless of the terms that their defenders use to characterize them, both sorts of social structures in the real world are a mish-mash of elements and have persistent problems. Those nations that are nearer the capitalist, "free trade" end of the spectrum find it impossible to eliminate the grinding

poverty of so many of their citizens, while people's republics tend to be dreadfully repressive. Perhaps scientists working in the "free world" are unconsciously being influenced by the character of their societies, but I for one prefer unconscious constraints to the explicit sort imposed by totalitarian states of all sorts. So do Levins and Lewontin, but they seem to think that the negative features of capitalist countries are inherent to them while the equally negative features of those societies that are trying to realize the Marxist ideal are only accidental and can be eliminated. I have my doubts.

Science developed in the West, and many of its characteristics may be merely remnants of its genesis. Levins and Lewontin sketch at least the general features of an alternative view of science. I happen to share many of their preferences but not their view of the historical contingencies that gave rise to them. □

David L. Hull is Professor in the Department of Philosophy, Northwestern University, Brentano Hall, Evanston, Illinois 60201, USA.

Geological history of the deep south

Jean-Claude Duplessy

South Atlantic Paleooceanography. Edited by K.J. Hsü and H.J. Weissert. Cambridge University Press: 1985. Pp.350. £40, \$69.50.

IN September, 1925, the Royal Research Ship *Discovery*, famous as Scott's first Antarctic vessel, sailed again for the South. Her purpose was not to explore new lands, but to make hydrological and biological observations in the South Atlantic by using "the same methods as those which have been already productive of valuable results in the North Atlantic". Fifty five years later, *Glomar Challenger* sailed to the South Atlantic and geologists might have used the same phrase in their proposal to study this basin. The geological history of the South Atlantic is still much more poorly known than that of the North Atlantic, which is the first to be studied by the most recent methods of geophysics and geochemistry.

It is not surprising, therefore, that the first drill cores obtained with *Glomar Challenger's* hydraulic piston corer during five cruises in the South Atlantic in 1980 have provided new details on the evolution of climatic conditions and ocean circulation and chemistry during the past 70 million years. These findings are summarized in *South Atlantic Paleooceanography*, which is based on 15 papers (including one extended abstract) presented at the First International Conference on

Paleoceanography held in Zürich in July 1983. The papers themselves have been edited thoroughly, and the text is clearly presented and illustrated. The overall structure, however, is a patchwork of individual studies dealing with one site or, at best, with the set of sites drilled during one leg, and no attempt has been made to compare results on the scale of the South Atlantic as a whole.

More than one-third of the book is devoted to sedimentological studies, which rely mainly on the carbonate, opal and organic carbon contents of the sediments. The carbonate content is a complex function of at least three parameters: carbonate production by plankton living in the surface waters; dissolution of carbonate shells during their fall to the sediment, or at the sediment-water interface; and the dilution of the remaining carbonate shells by detrital particles originating from the continents. The relative importance of these factors depends on the distance from the nearest continent, on the availability of nutrients supplied either by rivers or by upwelling, and on the global chemistry of the ocean. Nevertheless, geologists have demonstrated that there have been periods of intense calcite dissolution across the entire South Atlantic, a finding which constitutes a challenge for the modelling of the CO₂ cycle.

During the Cenozoic, the global cooling trend recognized in the oxygen isotope record is still dominant, but detailed analysis of some core sections of shorter duration shows a great deal of variability in the record, which re-opens the debate on the relative importance of the two dominant factors in this record, continental ice volume and sea water temperature. The occurrence of large carbon isotope shifts points to important changes in deep-water circulation, besides those affecting the carbon cycle, so that the interpretation of the oxygen isotope record might be a more complex matter than generally thought.

The last three chapters mainly discuss the non-isotopic evidence of deep-water variability, such as hiatuses, and sedimentological and micropalaeontological observations. They include various scenarios to explain the inception of the cold, deep-water realm about 40 million years ago and to account for the changes in deep-water circulation during the Neogene.

This book will be useful to marine geologists who want an up-to-date summary of some of the recent findings to emerge from the cruises of *Glomar Challenger* without having to search within those heavy blue books, the *Initial Reports of the Deep Sea Drilling Project*. □

Jean-Claude Duplessy is Director of the Centre des Faibles Radioactivités, Laboratoire Mixte CNRS - CEA, Parc du CNRS, 91190 Gif sur Yvette, France.

Pippard perturbed

Colin Gough

Response and Stability: An Introduction to the Physical Theory. By A.B. Pippard. Cambridge University Press: 1985. Pp.228. Hbk £27.50, \$49.50; pbk £9.95, \$17.95.

PHYSICS at an undergraduate level is almost exclusively concerned with linear processes. In contrast, the engineer and applied mathematician are taught from a very early stage of the real-life importance of non-linearity and of the often unexpected consequences of such behaviour. In this stimulating and thought-provoking textbook, Professor Pippard aims to broaden the background of the physicist into the traditional domain of the engineer. His discussion of the mathematics underlying phenomena as diverse as the vibration and collapse of loaded columns and the fluctuations and phase transitions of ferromagnets has resulted in a book of equal value to engineer and physicist. The author's view is all-embracing — even such unexpected topics as politics, economics, consciousness and religion get a brief mention.

The contents include a description of linear response theory, feedback, stability, non-linear systems, bifurcations, chaos and an introduction to catastrophe theory. Most of the first part of the book is concerned with engineering examples — the vibrations and stability of mechanical structures and electronic oscillators, the stability of water jets, convection currents and turbulence, and so on. In the last two chapters, on phase transitions and broken symmetry in solid-state physics, Professor Pippard extends the discussion of stability to include an introduction to Landau theory and fluctuations.

The book is designed both as a teaching

text and for individual study. Although the derivations of many important results are left as exercises for the reader, it is almost like having the author as a personal supervisor, such is the physical insight and guidance accompanying each of the problems; indeed, the book could be recommended as a lesson in problem-solving alone.

Professor Pippard is a great believer in practical demonstrations to illuminate "the physics" underlying the behaviour of mathematical equations. The book describes many simple experiments, which the reader is encouraged to repeat — incidentally providing a gold-mine of ideas for student projects. These include experiments with operational-amplifiers, the bending of flexible steel rules, the distortions of "slinky-type" springs and even the bouncing of a rubber ball between the walls of a shaken box. This latter problem is used as an example of a mechanical system whose transitions through a sequence of bifurcations to chaos can be analysed mathematically on a pocket calculator or microcomputer. Throughout the text, the numerical solution of mathematical models using a microcomputer is encouraged.

This is a book that every graduate physicist should read. No only does it provide a concise introduction to aspects of the real physical world with which experimental physicists should be familiar, but it also communicates the excitement and universality of recent research in areas as diverse as the physics of lasers and turbulence in liquids. The book should be of equal interest to the engineer, who will find the physical insight that Professor Pippard brings to many familiar problems a welcome change from the "take the Laplace transform" reflex of so many engineering textbooks. □

Colin Gough is a Senior Lecturer in the Department of Physics, University of Birmingham, PO Box 363, Birmingham B15 2TT, UK.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

"Life on the Prairie — The Buffalo Hunt" by Arthur F. Tait. (The hunter's musket was kept upright until the moment of firing to prevent the ball from rolling down the barrel.) The picture is taken from Heads, Hides & Horns: The Compleat Buffalo Book, by Larry Barsness, a thorough and well-illustrated history of the buffalo in North America. Publisher is Texas Christian University Press; price is hbk \$40, pbk \$19.50.

PLENUM: VISION IN RESEARCH

ADVANCED SCANNING ELECTRON MICROSCOPY AND X-RAY MICROANALYSIS

by Dale E. Newbury,
David C. Joy, Patrick
Echlin, Charles Fiori, and
Joseph I. Goldstein,

A graduate-level text detailing research topics at the leading edge of the field, including modeling electron beam-specimen interactions, SEM microcharacterization of semiconductors, electron channeling contrast, magnetic contrast, computer-aided imaging and interpretation, alternative microanalytical techniques, specimen coating, advances in specimen preparation for biological SEM, and cryo-microscopy.

0-306-42140-2/448 pp. + index/ill.
1986/\$37.50

(\$45.00 outside US & Canada)

ELECTRON ENERGY- LOSS SPECTROSCOPY in the Electron Microscope

by R. F. Egerton

Covers all the types of energy-loss measurements that can be carried out with an electron microscope, showing the experimentalist in detail how to maximize the performance of a spectrometer system and how to extract structural and chemical information from the energy-loss spectrum.

0-306-42158-5/410 pp. + index/ill.
1986/\$59.50

(\$71.40 outside US & Canada)

FOURIER TECHNIQUES AND APPLICATIONS

edited by John F. Price

A comprehensive overview for mathematicians, engineers, geologists, economists, and scientists working with either the theoretical or practical aspects of Fourier and spectral methods. Coverage begins with basic theory and extends to recent developments and new techniques and applications.

0-306-42100-3/proceedings/240 pp./ill.
1985/\$35.00

(\$42.00 outside US & Canada)

COME SEE PLENUM AT BOOTH 937
AT THE PITTSBURGH CONFERENCE!

Plenum Publishing Corporation
233 Spring Street
New York, N.Y. 10013

In the United Kingdom:
88/90 Middlesex Street
London E1 7EZ, England

Reader Service No.2

Plenum

Mixed tachykinins

Michael R. Hanley

Substance P: Metabolism and Biological Actions. Edited by C.C. Jordan and P. Oehme. Taylor & Francis: 1985. Pp. 260. £30, \$54.

OVER THE past two years, the substance P field has experienced an upheaval with the recognition that there is not one mammalian peptide, but rather a family of three or more substance-P-like "tachykinins". Moreover, there has been a gradual acceptance that pharmacological differences in the activities of tachykinins are attributable to a family of related receptors. Thus we have come to the sobering realization that much of the earlier work (of the past 50 years!) will have to be redone. This book, based on a symposium held in Maidstone, UK, in August 1984, is the first to gather the new ideas together. Consequently, it is disappointing that while the book as a whole is adequate it is hardly exciting.

The problems arise from two of the generic hazards of symposia volumes: widely different quality in the contributions and lack of a coherent style or format. For example, Chapter 1 is essentially an unrefereed journal article, whereas Chapter 7 is, at two-and-a-half pages, barely an extended abstract. The editors have elected to throw in the unedited abstracts of the meeting posters which, because of their brevity, are little more than souvenirs for the participants. Their inclusion is a distraction and the editors would have been better advised to expand the 14 contributed chapters.

Nevertheless, several of the contributions stand out and are worth special mention. Chapter 5 by Sandberg is excellent, bringing together a diffuse literature on the physical structure of substance P and relating it to its interaction with a single, precisely-defined target site (the smooth muscle of guinea pig ileum). Much is said and written about structure-function analysis, but rarely does the analysis go beyond interminable lists of structural analogues. Here, the clarity of the exposition and Sandberg's evaluation of the literature make the chapter both fun and illuminating.

Although "metabolism" appears in the book's title, only two chapters are devoted to it. Chapter 2 by Krause is the better of them. Beginning with a critical account of the concept of regulatory peptidases, the author proceeds to a generally useful dissection of the substance P literature concerned with peptidase action. This is a nice summary of an area that is difficult to follow in the primary literature. Chapter 10 by Lembeck focuses on the peripheral pharmacology of substance P and how it

relates to the effector actions of sensory nerves containing peptides. This chapter is a broad compendium of diverse actions, which are clearly explained, and culminates in a unique and thought-provoking summary of conditions or disorders in which tachykinins might be involved. Each of these three chapters is well crafted, unlikely to date quickly and accessible to a general audience. They contrast with most of the other contributions, which are rather pedestrian in style and content even though some of the work presented has been of great importance.

One aspect of the volume which is bound to irritate the general reader is the confusing nomenclature to tachykinins and their receptors. This is particularly ironic in as much as the book begins with a section on the topic! A standardized nomenclature has not yet been adopted,

but the introduction here of two new terms, "neurokinin A and B", for tachykinins named twice already doesn't help.

The final chapter, looking into future, encapsulates some of the drawbacks of the volume as a whole, in that it emphasizes certain aspects of the interest in substance P while neglecting others such as the genetics, receptor-linked biochemical events and regulation of cell growth, immune cell function and transporting epithelia. Indeed, in spite of the title, the emphasis of the book is on the pharmacology of substance P and its structural relatives. Overall, then, it can best be considered in the same way as a journal supplement, with its natural audience limited to the specialist. □

Michael R. Hanley is in the Medical Research Council's Molecular Neurobiology Unit, University of Cambridge Medical School, Hills Road, Cambridge CB2 2QH, UK.

Physics in fiction

Peter Kemp

Einstein as Myth and Muse. By Alan J. Friedman and Carol C. Donley. Cambridge University Press: 1985. Pp. 224. £25, \$39.50.

ALBERT Einstein was once invited to appear for a three-week season at the London Palladium. He declined. But, as Alan J. Friedman and Carol C. Donley's lively book reveals, he has turned up posthumously in equally unlikely contexts: on Chinese postage stamps, among the paranoid outpourings of lunatics, in advertisements for hosiery, business computers or beer. It is Einstein's infiltration of literature, though, that the authors are most interested in. Direct depictions of him are occasionally sighted, usually of a less than inspired kind: "Einstein has come among the daffodils/shouting/ that flowers and men/ were created/relatively equal", William Carlos Williams carols in one poem; in another Archibald MacLeish reports that "He is small and tight... And he terminates/In shoes". What the book mainly surveys, however, is the influence of Einstein's ideas upon literature.

In brisk preliminary chapters, the authors offer accounts of relativity theory and quantum physics that are concise and compulsive feats of intellectual exposition. Where things become less gripping is with the attempt to prove that the discoveries of modern physics have affected literature profoundly. One awkwardness Friedman and Donley keep encountering is that writers using or claiming to use Einstein's insights tend to be either obscure or obtuse. What is cited as "the most detailed, accurate, and knowledgeable parallel to the quantum theory that we have found in literature", Robert

Coover's novel *The Universal Baseball Association, Inc. J. Henry Waugh, Prop.*, is a name unlikely to leap instantly to many lips. Lawrence Durrell's *Alexandria Quartet* is more familiar: but, though Friedman and Donley pounce on the author's announcement that, with it, he's "trying to create a four-decker novel whose form is based on the relativity proposition", they ultimately have to admit that Durrell doesn't understand the difference between relativity theory and quantum theory, and harbours the elementary misconception that the "inescapable relativity of truth is an accurate reflection of the theory of relativity".

The authors are takingly keen to establish links between art and science, and are good at demonstrating the various impacts of Aristotle's and Newton's beliefs on literary imaginations. Their ingenious efforts to establish that Einstein's theories have likewise left their mark don't always carry the same degree of conviction, though. "Physicists found there was no universal frame of reference, and multiple viewpoints showed up... in novels", it is declared, implying some direct connection: yet multiple viewpoints were the speciality of the epistolary novel, in its heyday in the Newtonian eighteenth century. Quantum theory's revelation that "disintegration, violence, and derangement appear even on the sub-atomic level" could be responsible for contemporary literature's fascination with such phenomena, it is suggested — by-passing social, political and personal factors closer to hand and far more likely to have been influential. *Einstein as Myth and Muse* is an appealingly ambitious and stimulating study. Some of its musings, however, have a distinctly myth-like look themselves. □

Peter Kemp is an Associate Lecturer in the Department of English, Middlesex Polytechnic, All Saints, White Hart Lane, London N17 8HR, UK.

Evolution of the atmosphere and oceans

H. D. Holland, B. Lazar & M. McCaffrey

Department of Geological Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

The residence times of most constituents of the atmosphere and oceans are small fractions of the age of the Earth and, in general, their rate of output has been nearly equal to their rate of input. We are disturbing a number of these dynamic equilibria quite severely. The mineralogy of marine evaporites rules out drastic changes in the composition of sea water during the last 900 Myr. The chemistry of soils formed more than 1,000 Myr ago suggests that the atmosphere then contained significantly more CO₂ and less O₂ than at present. Hydrogen peroxide may well have been the principal oxidant and formaldehyde the main reductant in rain water between 3,000 and 1,000 Myr ago. Major changes in atmospheric chemistry since that time are almost certainly related to the evolution of the biosphere.

GEOCHEMISTS tend to regard the Earth as an extremely complicated but ultimately understandable chemical system. At present the system is nearly closed. Only small quantities of hydrogen and helium escape outwards into interplanetary space¹⁻³; cosmic rays, solar wind and a minor flux of meteorites annually contribute a small increment of mass to the Earth. The Earth is chemically very active, its chemistry being driven by both internal and external sources of energy. The internal sources derive largely from the decay of long-lived radionuclides, among which ²³⁸U, ²³⁵U, ²³²Th and ⁴⁰K are the most important^{4,5}. Energy released by these radionuclides drives convective motions in the Earth's mantle and possibly in the Earth's core. Mantle convection in turn gives rise to seafloor spreading, subduction, and mountain building, while core convection generates the Earth's magnetic field. Nearly all of the external energy that reaches the Earth is derived from the Sun. The energy flux from this source is about 5,000 times the internal energy flux and is responsible for the great intensity of the hydrological cycle, for the photochemistry of the atmosphere, and for green-plant photosynthesis.

Together the two energy sources drive the major geochemical cycles of the Earth. Mountain ranges raised by the supply of internal energy expose rocks to the levelling effects of physical erosion and chemical weathering. Most of the products of these processes are transported to the oceans. Solid residues accumulate as detrital marine sediments, largely within a few hundred kilometres of shorelines. Dissolved constituents in river water enter the oceans and are removed from sea water by a variety of reactions involving, among others, precipitation, ion exchange, and the alteration of ocean crust⁶.

The composition of the atmosphere and of sea water is determined by all these processes. The concentrations of most of the constituents of sea water exhibit very short response times, on a geological timescale. A few constituents show long response times, and the quantity of some is simply the sum of all of their inputs during Earth history. Most aspects of the chemistry of the atmosphere and of the oceans are now understood qualitatively^{7,8}, and a great deal of progress has been made recently in reconstructing their evolutionary history.

The ocean-atmosphere system

Reservoirs and residence times. The Earth is the only planet known to possess oceans, and the only planet with an atmosphere that contains large quantities of oxygen. These distinctions are, however, superficial. The oceans and the atmosphere contain very little of the mass of the Earth ($M = 6.0 \times 10^{24}$ kg), the mass of the oceans (1.4×10^{21} kg) amounting to only 0.0002 M

and the mass of the atmosphere (5.1×10^{18} kg) being somewhat less than $10^{-6} M$.

Both the oceans and the atmosphere can be represented as single reservoirs or boxes, with the movement of material into and out of these reservoirs represented by appropriate arrows. The resulting box models turn out to be useful in describing the operation of the Earth as a chemical system, in modelling the behaviour of its many constituents, and in predicting the response of the system to a variety of disturbances. Unfortunately, models in which a single box represents the world's oceans are frequently inadequate, because chemical heterogeneities within the oceans have an important role in the behaviour of many of the constituents of sea water⁹. To model the behaviour of these constituents we must subdivide the oceans into two or more reservoirs. This procedure improves the power and realism of the analyses; however, when the required number of reservoirs is large, the data needed to describe the behaviour of the system as a whole are often difficult to obtain.

The distribution of elements in the oceans is related in part to their residence times in the oceans. If we define a characteristic time τ_A for an element A such that

$$\tau_A = \frac{A}{\sum_i dA_i/dt}$$

where A is the total quantity of A in the oceans and $\sum_i dA_i/dt$ is the sum of the inputs of A into the oceans from all reservoirs, then we can distinguish two extreme cases: in the first the value of τ_A is much smaller than the age of the Earth; in the second it is roughly equal to the age of the Earth. The behaviour of calcium in the oceans illustrates the first case. As shown in Table 1, $\tau_{Ca^{2+}}$ for the oceans is just under 1 Myr; that is, about 1/5,000

Table 1 Concentration and mean residence time of the major constituents of sea water

Constituent	Average concentration in sea water of 35‰ salinity (mmol kg ⁻¹)	Residence time, τ (yr)
Na ⁺	468.0	4.8×10^7
Mg ²⁺	53.2	1.0×10^7
Ca ²⁺	10.2	8.5×10^5
K ⁺	10.2	5.9×10^6
Cl ⁻	545.0	7.3×10^7
SO ₄ ²⁻	28.2	7.9×10^6
HCO ₃ ⁻	2.4	8.0×10^4
Br ⁻	0.84	1.0×10^8

Data from ref. 7, p. 154; see also ref. 8, pp. 26, 27.

the age of the Earth. At the present rate of addition the concentration of Ca^{2+} would increase by a factor of 10 in ~ 8 Myr, unless the outputs of this element are a significant fraction of the inputs. All of the constituents of sea water have values of τ that are <100 Myr, that is, $<2\%$ of the age of the Earth. As the composition of sea water has not varied dramatically in the last 100 Myr (ref. 9, Ch. 9), the oceans must be a reservoir that solutes have both entered and left, and where the τ value of an element is a measure of its residence time in the oceans.

The τ value of the major constituents of sea water is considerably greater than the mixing time of the oceans ($\sim 1,000$ yr). These elements are therefore distributed quite uniformly throughout the oceans. Some minor constituents, Be, Fe and the rare earth elements among them, have residence times less than 1,000 yr; not surprisingly, these elements are distributed heterogeneously in the oceans. Many minor constituents of sea water whose τ values are well in excess of 1,000 yr are, however, also distributed heterogeneously. These include the elements that are actively involved in the biology of the oceans. Phosphorus⁸, for example, has a τ value of $\sim 7 \times 10^4$ yr. During photosynthesis this element is virtually completely stripped from surface waters, where it is usually a limiting nutrient; it then returns to sea water during the decomposition of organic matter. A good deal of particulate organic matter sinks and decomposes at depth. Deep ocean waters are therefore enriched in phosphorus relative to surface waters. Upwelling of these deep waters restores phosphorus to surface waters and ensures their continued fertility.

A few constituents of the atmosphere, such as Ne and Ar, have τ values which are not much shorter than the age of the Earth (ref. 9, Ch. 3). Degassing of the Earth contributes both elements to the atmosphere. They seem to be too heavy to escape from the atmosphere, and they are chemically too inert to be removed from the atmosphere into sediments. The quantity of Ne and Ar in the atmosphere is therefore almost certainly close to the sum of the quantity of each gas that was present early in Earth history plus the amount that has escaped since then from the solid Earth into the atmosphere¹⁰.

Inputs, outputs and control mechanisms. Much of the geochemical research which deals with the ocean-atmosphere system aims toward a quantitative formulation of the processes that determine the behaviour of its constituents. We are still some way from this goal, although a good deal of progress has been made, and the major sources and sinks of most of the cations and anions in sea water have now been identified (ref. 7, Ch. 5; ref. 11). We know, for instance, that river water is the major source of Ca^{2+} in the oceans, and that significant quantities of Ca^{2+} are also added to sea water by cation exchange reactions in estuaries¹² and by hydrothermal reactions during the cycling of sea water through mid-ocean ridges^{13,14}. The deposition of CaCO_3 is by far the most important output mechanism of Ca^{2+} from the present-day oceans, and the precipitation of gypsum ($\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$), anhydrite (CaSO_4) and dolomite ($\text{CaMg}(\text{CO}_3)_2$) plays a minor but significant part in the marine budget of Ca^{2+} . The kinetics of these removal processes are complicated. Surface sea water is supersaturated with respect to calcite and aragonite, the common polymorphs of CaCO_3 ⁸. Nevertheless, little CaCO_3 is precipitated inorganically; virtually all CaCO_3 is precipitated biologically in near-surface waters. Much of this material sinks and is redissolved in the deeper parts of the oceans, where sea water is undersaturated with respect to both calcite and aragonite (ref. 8, Ch. 2). The shift from supersaturation to undersaturation is a consequence of the progressive increase in total pressure, the decrease in temperature, and the increase in the CO_2 content of sea water with depth. The third of these factors is related to the oxidation of organic matter in the deeper parts of the oceans.

A quantitative formulation of the relationship between the state of the oceans and the rate of removal of Ca^{2+} as a constituent of CaCO_3 demands a thorough understanding of the

rate of CaCO_3 production in the surface oceans, the change with depth in the degree of saturation of the oceans with respect to calcite and aragonite, the kinetics of the dissolution of these phases, and the rate of return of Ca^{2+} from deep- to near-surface portions of the oceans. Several parts of this complex system are now reasonably well understood, but current estimates of the response of the system to changes in the physics, chemistry and biology of the oceans are still quite imprecise. This is equally true for predictions regarding the behaviour of the other major constituents of sea water, and it will probably require another decade or two of intensive study before we have a reasonably quantitative model of the relationship between the concentration of these constituents in sea water and their rate of output from the oceans.

Anthropogenic and non-anthropogenic influences on P_{CO_2} . CO_2 is probably the most intensively studied constituent of the atmosphere-ocean system. Variations in the concentration of CO_2 in the atmosphere have been monitored carefully since 1958. The relationship between the observed increase in P_{CO_2} , the rate of fossil fuel burning, and changes in the mass of the Earth's biosphere has been analysed thoroughly and much debated^{15,16}. The role of changes in the biosphere is still uncertain, because it is difficult to measure gains and losses in the mass of the biosphere with sufficient precision¹⁷. Models of the downward mixing of CO_2 in the oceans have been improved considerably, but the uncertainty in figures for the downward mixing rate is still approximately 20%. It is clear, however, that the present rate of CO_2 addition to the atmosphere due to fossil fuel burning exceeds all of the non-anthropogenic fluxes of CO_2 except those of photosynthesis and decay¹⁷. We are seriously disturbing the operation of the carbon cycle; significant changes in the Earth's climate may result. The decrease of fossil fuel reserves will probably lead to a diminished rate of CO_2 addition to the atmosphere within the next 150 yr, but the removal of the excess CO_2 from the atmosphere will probably require on the order of 10,000 yr.

Non-anthropogenic changes in the CO_2 content of the atmosphere during the last ice age have been documented recently by studies of air trapped in ice cores from Greenland and Antarctica^{18,19}. During the peak of the last ice age the CO_2 content of the atmosphere was ~ 180 p.p.m. During the nineteenth century it was probably close to 280 p.p.m.; its present value is 345 p.p.m. Neither the cause of the low CO_2 content of the atmosphere during the height of the last glaciation nor its climatic effects are presently well understood²⁰⁻²³.

It is likely that the CO_2 content of the atmosphere has varied even more widely on the much longer timescale of Earth history. During the past several years the first serious attempts have been made to relate P_{CO_2} to changes in climate during the past 100 Myr²⁴⁻²⁶. The results are intriguing, but the complexity of the system and the number of parameters that affect climate are so large that current models should probably be treated with caution. The behaviour of other constituents of the ocean-atmosphere system is no easier to model. Their concentration could and almost certainly has varied significantly during Earth's history, as the response time of nearly all components of this system is much less than the age of the Earth. In the absence of convincing models that can be used to reconstruct the compositional evolution of the atmosphere and oceans, we must look to the geological record for direct indicators of the physical state and the chemical composition of the ocean-atmosphere system in the past.

The ocean-atmosphere system in the past

The major components of sea water. In 1961 L. G. Sillén proposed an equilibrium model for the chemistry of oceans²⁷, in which he suggested that all of the major compositional parameters of sea water are controlled by thermodynamic equilibrium between dissolved species in sea water and mineral phases in marine sediments. The implications of the model were profound. If it

had proved to be correct, the chemical composition of sea water would have been essentially locked into its present state during much of Earth history.

The model did not survive the detailed scrutiny to which it was subjected during the decade after its publication. The formation of magnesian clay minerals in the oceans, which the model had predicted, was not observed²⁸. Instead, sea water cycling through mid-ocean ridges emerged as a major process for removing magnesium from the oceans^{13,14,29}. A careful look at the mechanisms of the removal of SO_4^{2-} from the oceans showed that processes such as bioturbation, SO_4^{2-} diffusion into sediments^{30,31}, and evaporite basin development³² rather than thermodynamics control the output of this ion from the oceans, and that the nature and intensity of the processes that have determined the kinetics of the dominant output processes of this and other ions have probably varied significantly during Earth history. At the same time, it was observed that the sequence of mineral phases formed in evaporite basins by the progressive evaporation of sea water has not varied significantly during the 900 Myr for which a record of marine evaporites is known to exist³³.

Students of marine evaporites have therefore tended to assume that the composition of sea water has been constant during this part of Earth history³⁴; this assumption is not necessarily valid, as small variations in the composition of sea water do not necessarily alter the sequence of minerals formed during the evaporation of sea water. What must be established are the limits to changes in the composition of sea water imposed by the condition that the deposition sequence of evaporite minerals remains invariant. This problem was tackled more than a decade ago³⁵, and its solution has been refined recently (ref. 9, Ch. 9; refs 36–38). It turns out that the concentration of nearly all of the major constituents of sea water could have varied by a factor of 2 to 3 in either direction without a change in the early part of the normal mineral sequence in marine evaporites, the part for which the best historical evidence is available. This result is interesting and intuitively reasonable; it also demonstrates that criteria other than the mineralogy of marine evaporites are needed to define with some precision the evolution of sea water composition during the past 900 Myr.

Fluid inclusions in halite (NaCl) from marine evaporites are a potential source of this information. During halite precipitation small amounts of brine are usually trapped in NaCl crystals. We have observed fluid inclusions in halite crystals from the salt pans on Great Inagua Island (Bahamas) and similar inclusions are almost universal in halite of all ages (ref. 39, Ch. 11). Recently, fluid inclusions in halite have attained notoriety as a possible source of danger for canisters of high-level nuclear wastes stored in repositories that are sited in salt⁴⁰. The fate of brines trapped in halite is not well known. Some brine inclusions have obviously moved significant distances during the recrystallization of the enclosing salt. Some were introduced well after the halite had crystallized, and may have no genetic connection with the brines from which the halite was originally precipitated. On the other hand, inclusion fluids in salt which has suffered little from recrystallization and from later brine invasion may well be representative of the brines from which the halite precipitated. The pioneering work of Holser⁴¹ suggested that such fluid inclusions do indeed exist, and our own work confirms his findings.

During the evaporation of sea water today the precipitation of a small amount of CaCO_3 with/without $\text{CaMg}(\text{CO}_3)_2$ is followed by the precipitation of large quantities of CaSO_4 with/without $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$. Halite begins to precipitate when ~90% of the initial amount of water has been lost by evaporation. Prior to halite saturation, most of the HCO_3^- and Ca^{2+} in sea water, and approximately one-third of the SO_4^{2-} , are lost to precipitates; Na^+ , K^+ , Mg^{2+} and Cl^- are essentially conserved. After the onset of NaCl precipitation, Na^+ and Cl^- begin to be lost from solution; the behaviour of these ions in the residual

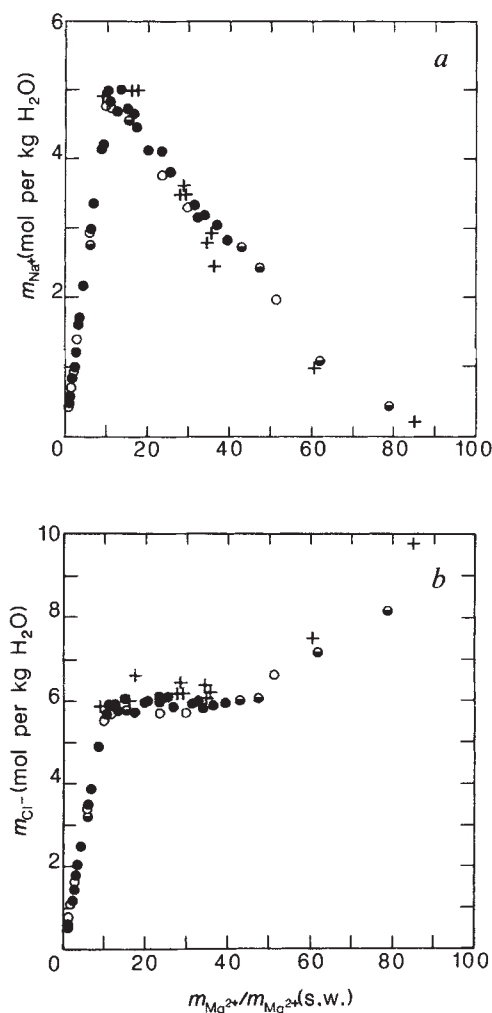


Fig. 1 The evolution of the concentration of a, sodium and b, chloride ions during the evaporation of present-day and Miocene sea water, plotted against a measure of the degree of evaporation. ●, Salinas on Great Inagua Island; ○, taken from ref. 43; ◐, from ref. 44; +, fluid inclusions in Miocene halite from the Red Sea.

brines reflects the value of the Na^+/Cl^- ratio in sea water. Mg^{2+} and K^+ are largely conserved during halite precipitation until the latest stages of evaporation, during which sulphates and chlorides of Mg^{2+} and K^+ precipitate^{33,34}.

In Fig. 1 the concentrations of Na^+ and Cl^- in evaporating sea water are plotted against the ratio of the concentration of Mg^{2+} in the brines to the concentration of Mg^{2+} in sea water. The data for the evaporation path of present-day sea water derive from our study of the salinas on Great Inagua Island⁴² and from two laboratory studies^{43,44}. The concentration of both ions rises conservatively until the beginning of halite precipitation; thereafter the concentration of Na^+ decreases steeply, as the ratio $m_{\text{Na}^+}/m_{\text{Cl}^-}$ in sea water today is less than unity.

The crosses in Fig. 1 indicate the relationship between the Na^+ , Cl^- and Mg^{2+} concentration of inclusion fluids extracted from samples of Miocene halite in the DSDP core at site 227 in the Red Sea⁴⁵. Holes were drilled into these inclusions with a microdrill, fluid was extracted with a micropipette, the volume of the extracted fluid was measured by determining the length of a capillary of known cross-sectional area which was occupied by the inclusion fluid, and the composition of the brines was determined by means of a Dionex ion chromatograph. The similarity of the evaporation path of modern sea water and the evolution of Miocene sea water as recorded by the composition

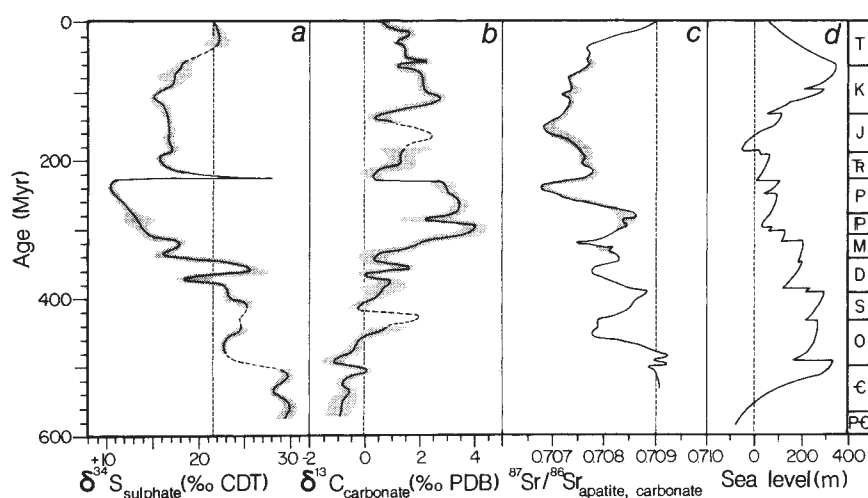


Fig. 2 The time variation of *a*, sulphur isotopes in evaporite sulphate; *b*, carbon isotopes in marine carbonates; *c*, strontium isotopes in marine carbonates and in fossil apatite; and *d*, sea level during the last 600 Myr⁶². PC, Precambrian; C, Cambrian; O, Ordovician; S, Silurian; D, Devonian; M, Mississippian; P, Pennsylvanian; P, Permian; R, Triassic; J, Jurassic; K, Cretaceous; T, Tertiary.

of the inclusion fluids is striking. This was expected, because the 5–6 Myr that have elapsed since the formation of these Miocene halites is less than the oceanic residence time of Mg^{2+} (~10 Myr) and much less than the oceanic residence time of Na^+ (~48 Myr) and Cl^- (~73 Myr). It would therefore have been surprising if the relationship between the concentration of the three ions in Miocene sea water had differed greatly from that in present-day sea water. The happy surprise was to find that there has been little or no alteration of this relationship within the inclusions in the Miocene halite. The concentration of the other major constituents in these fluid inclusions is not the same as that in the original brines, but the differences can be understood quite reasonably in terms of their interaction during recrystallization with halite which was located within a radius of a few millimetres of each brine inclusion⁴⁶. Much the same is true for inclusion fluids in halite from the Upper Permian Salado Formation near Carlsbad, New Mexico, and from the Lower Permian Wellington Formation near Lyons, Kansas⁴⁷. However, the Mg^{2+} and SO_4^{2-} concentrations of the brines in several of these inclusions have been depleted by dolomitization followed by anhydrite or gypsum precipitation. This process has gone very far in three fluid inclusions from Silurian halite in the Michigan Basin. The reconstruction of the composition of sea water from the composition of the Silurian inclusions is therefore a rather uncertain process. The evidence to date indicates strongly that the concentration of the major ions in Permian sea water was very close to their concentration in present-day sea water; the data indicate less strongly that this similarity extends to sea water during the Silurian. These conclusions corroborate those of Harvie *et al.*³⁸, who found that the mineralogy of the Permian Zechstein evaporites can be explained quite readily in terms of the evaporation of present-day sea water.

The isotopic composition of solutes in sea water. The available fluid inclusion data are obviously still very fragmentary. However, the study of the isotopic composition of sulphur^{48–50} and oxygen⁵⁰ in anhydrite from marine evaporites, and studies of the isotopic composition of carbon^{51–54} and strontium^{55–61} in marine limestones, cover a sizeable fraction of Earth history with a fair degree of completeness. All of these studies suffer from uncertainties regarding the purely marine origin of their contained elements and from the possibility that post-depositional processes have affected the isotopic composition of their constituents. However, as shown in Fig. 2, the patterns of fluctuations in the isotopic composition of sulphur and oxygen in seawater sulphate, of carbon in seawater bicarbonate, and of seawater Sr^{2+} have emerged very clearly⁶². The variations are all easily measurable, and those of $\delta^{34}\text{S}$ are large (+10% to +30% relative to the Canyon Diablo Meteorite Standard).

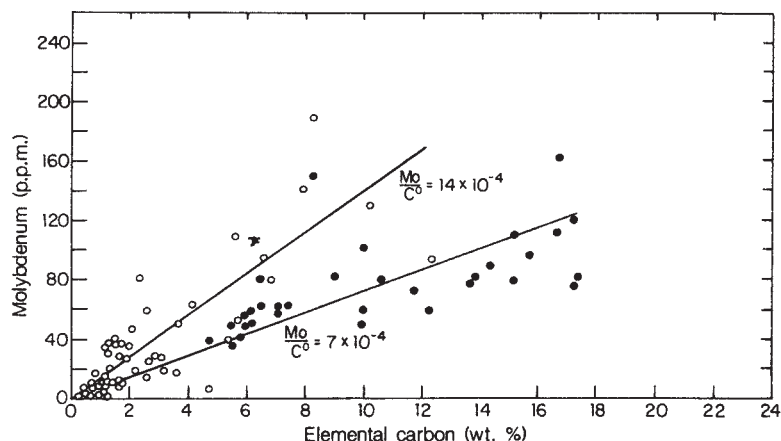
Changes in $\delta^{13}\text{C}_{\text{HCO}_3^-}$ are frequently coupled to changes in $\delta^{34}\text{S}_{\text{SO}_4^{2-}}$. During periods of very positive $\delta^{13}\text{C}_{\text{HCO}_3^-}$, $\delta^{34}\text{S}_{\text{SO}_4^{2-}}$ was most negative, and vice versa. This relationship is best interpreted in terms of the requirement that the rate of use of atmospheric oxygen during the oxidative weathering of organic matter to CO_2 and of reduced sulphur to sulphate be nearly balanced by the recovery of oxygen during the photosynthetic reduction of CO_2 followed by the burial of organic matter, and by the reduction of SO_4^{2-} to sulphide followed by the burial of sulphide minerals^{63–66}. The mechanisms by which the carbon cycle is coupled to the sulphur cycle are complex and not very well understood.

The fluctuations in the isotopic composition of Sr in sea water are largely due to changes in the isotopic composition of this element in average river water and to changes in the rate of cycling of sea water through mid-ocean ridges, where isotopically heavy sea water is exchanged for isotopically lighter Sr in mid-ocean ridge basalts^{67,68}. The relative importance of the two processes is still uncertain. It may be possible to resolve the matter by determining the evolution of the mean isotopic composition of Nd in sea water^{69,70}, as the isotopic composition of this element seems to be determined largely by its composition in river waters and only slightly by the cycling of sea water through mid-ocean ridges.

Trace elements in sea water. For several decades after the major-element chemistry of sea water had become well known, the concentration of many of the minor constituents of sea water was still uncertain by several orders of magnitude. During the past 20 years improvements in analytical techniques, greater care to avoid contamination, and increases in the size of samples taken for analysis have led to a reasonably complete knowledge of the distribution of most of the minor elements in the oceans^{71,72}. Our knowledge of the removal mechanism of many of these elements is still fragmentary, but it is well known that many of the metals are removed very efficiently from sea water into sediments that are rich in organic matter^{73,74}. Molybdenum is a case in point. Figure 3 shows the correlation of the molybdenum with the organic carbon content of sediments from the Black Sea. The mean Mo/C^0 ratio of the two sets of data is $\sim 10 \times 10^{-4}$ g per g. This ratio seems to have changed surprisingly little during much of Earth history⁷⁴. Figure 4, for example, illustrates the correlation of the molybdenum content and the concentration of organic carbon in Devonian black shales from the eastern United States. The ratio of Mo to C^0 is 12×10^{-4} , that is, very close to the value of this ratio in Black Sea sediments. This similarity extends back in time at least 2,000 Myr to the black schists of the Outokumpu⁷⁵ and Talvivaara regions in Finland (X. Ervamaa, personal communication).

The behaviour of uranium in black shales is similar to that

Fig. 3 The correlation between the molybdenum and the organic carbon content of sediments from the Black Sea. Data from refs 118 (○) and 119 (●).



of molybdenum. Uranium and organic carbon are well correlated in marine sediments of all ages. In Devonian black shales from the eastern United States the U/C^0 ratio is $\sim 4 \times 10^{-4}$ g per g.⁷⁶ In the Proterozoic black shales of the Outokumpu and the Talvivaara regions in Finland the U/C^0 ranges from $\sim 2.5 \times 10^{-4}$ g per g to 10×10^{-4} g per g (ref. 75 and Ervamaa, personal communication). This implies that enough oxygen or other oxidants were present in the atmosphere some 2,000 Myr ago to ensure the oxidation and dissolution of much of the uranium present in rocks undergoing weathering; it also implies that the removal mechanisms of uranium from the oceans have changed little since that time. On the other hand, the survival of grains of uraninite during weathering, and their concentration into large uranium ore deposits more than $\sim 2,000$ Myr ago⁷⁷⁻⁸⁹, suggests that the oxygen content of the atmosphere was then considerably less than it is now.⁸³⁻⁹⁰ This observation leads quite directly to the search for oxygen barometers that can be used to determine P_{O_2} in the past.

The history of atmospheric oxygen. Molecular oxygen is produced at a very rapid rate by green-plant photosynthesis. If none of the O_2 produced in this manner were removed by respiration and decay, the O_2 level in the atmosphere would double in approximately 10^4 yr. Virtually all of the O_2 produced photosynthetically in a given year is, of course, lost within a few decades by the oxidation of organic matter. Only about 0.2% of the organic matter generated via photosynthesis escapes re-oxidation to CO_2 and most of this small remainder is buried with marine sediments (ref. 7, Ch. 5). Marine sediments also contain reduced sulphur, most of which is a constituent of pyrite, FeS_2 . It can be shown (ref. 7, Ch. 6) that the generation of reduced carbon from CO_2 and of reduced sulphur from SO_4^{2-} frees O_2 at such a rate that the O_2 content of the atmosphere could double in a few million years. Such a doubling is prevented in large part by the removal of molecular O_2 by the oxidation of elemental carbon and pyrite in rocks undergoing weathering at the Earth's surface, and by the oxidation of the reduced components (largely H_2 , CO and SO_2) of volcanic gases. The O_2 budget of the atmosphere is apparently nearly balanced today, and seems to be determined by the requirement that the quantity of reduced carbon and sulphur in new sediments is nearly equal to that in sediments undergoing weathering. The relatively high oxygen content of the present atmosphere is demanded by the requirement that 99.8% of the huge amount of organic matter that is generated by green-plant photosynthesis be oxidatively destroyed. The functional relationship between P_{O_2} and the burial rate of organic carbon in marine sediments is not completely understood. The connecting parameters include the rate of photosynthesis on land and at sea, the O_2 content of sea water, the mixing time of the oceans, the distribution of bacteria in the oceans and marine sediments, and the behaviour of

nutrients—particularly phosphorus—in the oceans (ref. 8, Chs 3, 6; ref. 9, Ch. 8).

P_{O_2} could have varied significantly during the past few hundred million years. However, the continuity of animal life indicates that P_{O_2} has not fluctuated wildly. P_{O_2} was almost certainly never less than 10% nor more than five times its present level since the start of the Cambrian some 570 Myr ago (ref. 9, Ch. 9). The value of P_{O_2} during the Precambrian is only poorly constrained by our knowledge of the Precambrian biota, and a variety of highly divergent opinions have been expressed regarding the course of P_{O_2} during the first 85% of Earth history (ref. 9, Chs 7, 8; refs 91-104). The diversity of these opinions is a fair measure of the uncertainty attached to the criteria on which they were based. Perhaps the most promising new approach to setting at least broad limits on the O_2 content of the Precambrian atmosphere is the study of Precambrian soils. Palaeosols from this period are relatively rare, but a fair number are now known, and the active search for additional palaeosols will almost certainly reveal many others. The Precambrian palaeosols which have been studied in some detail were developed on rocks of either roughly granitic or roughly basaltic composition (ref. 9, Ch. 7). In all of the palaeosols developed on granitic rocks, iron was conserved during weathering, and is now present as a constituent of ferric oxides. This situation is similar to that in highly oxidized soils today. On the other hand, iron appears to have been lost from the upper layers of all of the palaeosols developed on fresh basaltic rocks; this behaviour indicates that

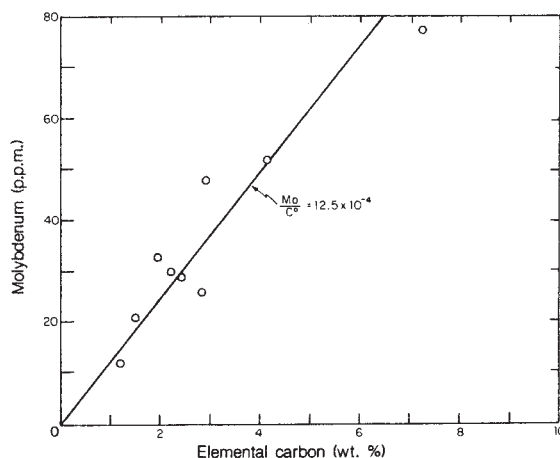


Fig. 4 The correlation between the concentration of molybdenum and organic carbon in Devonian black shales from the eastern United States⁷⁶.

dissolved oxidants became depleted in these soils, that Fe^{2+} released from ferromagnesian minerals was not oxidized to Fe^{3+} , and that iron was therefore transported out of these soils rather than retained in ferric oxide and/or hydroxide precipitates. The difference in the behaviour of iron in these two types of soils can be used to set broad limits on the ratio of P_{O_2} to P_{CO_2} in the Precambrian atmosphere¹⁰⁴.

The behaviour of Fe^{2+} and of other cations which can be oxidized during soil formation from a particular parent rock depends on the relative rate of supply of oxidants and weathering acids. Quantitatively, carbonic acid has always been the dominant weathering acid (ref. 9, Ch. 5), although organic acids may play a significant part in weathering today¹⁰⁵. Molecular O_2 is the major oxidant today. Oxidants such as H_2O_2 are produced photochemically in the atmosphere, but their concentration in rain water is usually less than 10% of the concentration of molecular O_2 (ref. 106). This may not have been the case during much of pre-Silurian time. The production rate of H_2O_2 and other oxidants in the atmosphere is only a weak function of P_{O_2} ; hence, the ratio of the concentration of photochemically produced oxidants to dissolved molecular O_2 in rain water increases rapidly with decreasing P_{O_2} , and probably exceeds unity at values of P_{O_2} between 2×10^{-3} and 2×10^{-4} atm¹⁰⁷. In rain water falling through such a low- O_2 atmosphere, the oxidant concentration therefore reflects more the photochemistry of the atmosphere than its O_2 content. It can be shown, however, that the rate of supply of O_2 to soils by diffusive transfer from the atmosphere is much more rapid than the rate of supply of oxidants dissolved in rain water, except at extremely low values of P_{O_2} and in abnormally thick soils¹⁰⁴. Hence the supply of oxidants appears to have been determined largely by the O_2 gradient and by the diffusion constant of O_2 in soils.

Today, CO_2 is supplied to soils largely by the bacterial oxidation of organic matter in soils¹⁰⁸⁻¹¹⁶. Before the development of higher land plants, CO_2 was probably supplied in large part as a dissolved constituent in rain water and by gaseous diffusion from the atmosphere. The solubility of CO_2 in water is much greater than that of O_2 ; hence, dissolved CO_2 played a larger role than dissolved O_2 in the budget of these gases in pre-Silurian soils. It seems likely, however, that the ratio of the rate of supply of O_2 to that of CO_2 in pre-Silurian soils was roughly proportional to the ratio of the gradient of O_2 to that of CO_2 between the atmosphere and the water table¹⁰⁴.

Carbonic acid reacts with a wide variety of minerals and releases cations, the charge of which is balanced by HCO_3^- . Among the major cations released during these reactions, Fe^{2+} is usually the most abundant reduced species. In the absence of organic matter and sulphide minerals, Fe^{2+} is therefore usually the most important potential sink for oxidants in soil waters. If the oxidant demand for the oxidation of Fe^{2+} is greater than the oxidant supply in soil water, not all of the Fe^{2+} released by soil acids will be oxidized within the soil; iron is then incompletely retained in the soil as a constituent of precipitated ferric oxides and/or hydroxides. It can be shown that Fe^{2+} can be completely oxidized when the ratio of the supply of oxidants to the supply of acids in soils is greater than R , where

$$R = \frac{M_{\text{FeO}}}{8(M_{\text{CaO}} + M_{\text{MgO}} + M_{\text{Na}_2\text{O}} + M_{\text{K}_2\text{O}})}$$

and where M_i = moles of constituent i per kg of parent rock.

The value of R is considerably greater for basalts than for granites. A given atmosphere can therefore contain an O_2 concentration which is sufficiently large for all of the Fe^{2+} released during the weathering of granites to be oxidized and retained as $\text{Fe}(\text{OH})_3$, but which is insufficient for all of the Fe^{2+} released during the weathering of basalts to be oxidized and retained.

The presently available evidence for the behaviour of iron in Proterozoic palaeosols suggests that $P_{\text{O}_2}^0$, the O_2 pressure in the atmosphere, was related to $P_{\text{CO}_2}^0$, the CO_2 pressure in the atmos-

phere, and to $P_{\text{CO}_2}^L$, the mean value of the CO_2 pressure at the water table, by the expression

$$P_{\text{O}_2}^0 \approx 0.03(P_{\text{CO}_2}^0 - P_{\text{CO}_2}^L)$$

The extraction of a value for $P_{\text{O}_2}^0$ requires an estimate of the CO_2 gradient in soils.

The history of atmospheric CO_2 . The residence time of CO_2 in the present-day atmosphere is only about 10 years with respect to cycling through the biosphere. Even the residence time of carbon in the sum of all the 'exchangeable reservoirs', that is, the atmosphere, the biosphere, and the oceans, is only 10^5 yr, and it is clear that the CO_2 content of the atmosphere is determined by the requirement that the output of carbon as a constituent of carbonates and organic matter in sediments be nearly equal to the rate of input of carbon compounds into the atmosphere from volcanoes, hydrothermal vents and weathering processes (ref. 7, Ch. 6). These output rates depend in a rather complicated manner on a host of variables that include the Earth's climate, the nature of the biosphere, the disposition of the oceans and continents, and the circulation pattern of the oceans. It is difficult, therefore, to reconstruct the evolution of P_{CO_2} from the nature of the geological record, but a first attempt can be made.

A study of thousands of chemical analyses of sedimentary rocks has shown that the degree of acid titration of igneous rocks that have undergone conversion to sedimentary rocks has varied little during the last 3,000 Myr (ref. 9, Ch. 5). Nearly all of the CaO and roughly half of the MgO in igneous rocks has always been converted into carbonates in sedimentary rocks. As the ratio of organic carbon to carbonate carbon in sedimentary rocks has also varied rather little, the CO_2 content of the atmosphere must have adjusted itself in geologically very short periods of time, so that the degree of carbonate titration of rocks undergoing weathering has remained reasonably constant during much of Earth history. Among the many changes in the atmosphere-biosphere-ocean-crust system that have affected P_{CO_2} , the development of a cover of higher land plants and the progressive increase in solar luminosity may well have been the most important¹¹⁷. Both changes have tended to lower atmospheric P_{CO_2} during the course of geologic history. The intensity of chemical weathering today is due in large part to the high CO_2 content of soil air, which is typically 10 to 100 times higher than the CO_2 content of the atmosphere¹⁰⁸⁻¹¹⁶. Prior to the existence of a cover of higher land plants, the difference between the CO_2 content of the atmosphere and soil air was almost certainly much smaller than it is today. If the proper rate of CO_2 output from the atmosphere required a level of CO_2 in soil air that is comparable to present values, then P_{CO_2} in the atmosphere must have been significantly higher before the development of a cover of higher land plants. A decrease in atmospheric P_{CO_2} following the advent of land plants would also have affected the Earth's climate and hence the distribution of the biota. The calculation of the decrease in P_{CO_2} that followed the development of higher land plants is therefore not straightforward.

The mean surface temperature of the Earth would decrease quite dramatically if the solar luminosity were to decrease significantly today. Since there has apparently been no long-term trend in the mean surface temperature of the Earth during the last ~2,000 Myr, the increasing solar luminosity during this interval must have been offset by a decrease in the intensity of the combined greenhouse effect of atmospheric gases. Changes in the partial pressure of several greenhouse gases, either singly or in combination, could have been responsible for this decrease¹¹⁷. A progressive decrease in P_{CO_2} alone could have accounted for the entire effect. The magnitude of the required decrease is hard to calculate with confidence, since even the temperature increase due to a doubling of the CO_2 content of the present atmosphere is known with a precision of no better than a factor of two.

All that can be said with confidence is that P_{CO_2} in the atmosphere was considerably higher during the Precambrian than it is today. A difference of a factor of 10–30 between P_{CO_2} during mid-Precambrian time and the present is likely¹¹⁷ but uncertain. A combination of this estimate with the available palaeosol data indicates that at that time P_{O_2} was on the order of 3×10^{-5} atm. The uncertainty in this estimate is probably an order of magnitude, but could be larger¹⁰⁴. Despite the very large uncertainty in this estimate, it is likely that free oxygen was present in the atmosphere during mid-Proterozoic time, but

that P_{O_2} was very much smaller than it is today. The behaviour of iron in palaeosols should be traced forward in time to determine when this element began to be retained in soils developed on basalts as well as in soils developed on granites. The behaviour of iron in palaeosols should also be traced backwards in time to determine when iron was retained neither in soils developed on granites nor in soils developed on basaltic rocks. It would be fascinating if these periods of change in the behaviour of iron in palaeosols coincided with evolutionary changes in the nature of the biota.

1. Hunten, D. M. & Strobel, D. F. *J. atmos. Sci.* **31**, 305–317 (1973).
2. Hunten, D. M. & Donahue, T. M. *A. Rev. Earth planet. Sci.* **4**, 265–292 (1976).
3. Torr, D. G. in *The Photochemistry of the Atmosphere* (ed. Levine, J. S.) Ch. 6 (Academic, New York, 1985).
4. Turcotte, D. L. *Earth planet. Sci. Lett.* **48**, 53–58 (1980).
5. O'Nions, R. K. & Oxburgh, E. R. *Nature* **306**, 429–431 (1983).
6. Garrels, R. M. & Mackenzie, F. T. *Evolution of Sedimentary Rocks* (Norton, New York, 1971).
7. Holland, H. D. *The Chemistry of the Atmosphere and Oceans* (Wiley, New York, 1978).
8. Broecker, W. S. & Peng, T.-H. *Tracers in the Sea* (Lamont-Doherty Geological Observatory, 1982).
9. Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans* (Princeton University Press, 1984).
10. Cogley, J. G. & Henderson-Sellers, A. *Rev. Geophys. Space Phys.* **22**, 131–175 (1984).
11. Mackenzie, F. T. & Garrels, R. M. *Am. J. Sci.* **264**, 507–525 (1966).
12. Sayles, F. L. & Mangelsdorf, P. C. *Jr. Geochim. cosmochim. Acta* **41**, 951–960 (1977).
13. Edmond, J. M. *et al. Earth planet. Sci. Lett.* **46**, 1–18 (1979).
14. Mottl, M. J. *Bull. geol. Soc. Am.* **94**, 161–180 (1983).
15. *Changing Climate, Report of the Carbon Dioxide Assessment Committee* (National Academy Press, Washington DC, 1983).
16. *Potential Climatic Impacts of Increasing Atmospheric CO₂ with Emphasis on Water Availability and Hydrology in the United States* (U.S. Environment Protection Agency, 1984).
17. Sundquist, E. T. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archaean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 5–60 (American Geophysical Union, Washington DC, 1985).
18. Neftel, A., Oeschger, H., Schwander, J., Stauffer, B. & Zundbrunn, R. *Nature* **295**, 220–223 (1982).
19. Delmas, R. J., Ascencio, J.-M. & Legrand, M. *Nature* **284**, 155–157 (1980).
20. Broecker, W. S. *Geochim. cosmochim. Acta* **46**, 1689–1705 (1982).
21. Ennever, F. K. & McElroy, M. B. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archaean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 154–162 (American Geophysical Union, Washington DC, 1985).
22. Toggweiler, J. R. & Sarmiento, J. L. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archaean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 163–184 (American Geophysical Union, Washington DC, 1985).
23. Wenk, T. & Siegenthaler, U. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archaean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 185–194 (American Geophysical Union, Washington DC, 1985).
24. Berner, R. A., Lasaga, A. C. & Garrels, R. M. *Am. J. Sci.* **283**, 641–683 (1983).
25. Berner, R. A. & Barron, E. J. *Am. J. Sci.* **284**, 1183–1192 (1984).
26. Lasaga, A. C., Berner, R. A. & Garrels, R. M. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archaean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 397–411 (American Geophysical Union, Washington DC, 1985).
27. Sillén, L. G. in *Oceanography* (ed. Sears, M.) 549–581 (American Association for the Advancement of Science, Washington, 1961).
28. Drever, J. I. in *The Sea Vol. 5* (ed. Goldberg, E. D.) Ch. 10 (Wiley, New York, 1974).
29. Thompson, G. in *Hydrothermal Processes at Seafloor Spreading Centers* (eds Rona, P. A., Boström, K., Laubier, L. & Smith, K. L.) 225–278 (Plenum, New York, 1983).
30. Berner, R. A. *Phil. Trans. R. Soc. A315*, 25–38 (1985).
31. Berner, R. A. *Geochim. cosmochim. Acta* **48**, 605–615 (1984).
32. Zharkov, M. A. *History of Paleozoic Salt Accumulation* (transl. Sorkina, R. E., Fursenko, R. V. & Vasilieva, T. I.) (Springer, New York, 1981).
33. Borchert, H. & Muir, R. O. *Salt Deposits; The Origin, Metamorphism, and Deformation of Evaporites* (Van Nostrand, New York, 1964).
34. Braitsch, O. *Entstehung und Stoffbestand der Salzlagertstätten* (Springer, Berlin, 1962).
35. Holland, H. D. *Geochim. cosmochim. Acta* **36**, 637–651 (1972).
36. Eugster, H. P. & Jones, B. F. *Am. J. Sci.* **279**, 317–334 (1979).
37. Eugster, H. P., Harvie, C. E. & Weare, J. H. *Geochim. cosmochim. Acta* **44**, 1335–1347 (1980).
38. Harvie, C. E., Weare, J. H., Hardie, L. A. & Eugster, H. P. *Science* **208**, 498–500 (1980).
39. Roedder, E. in *Reviews of Mineralogy Vol. 12* (ed. Ribbe, P. H.) (Mineralogical Society of America, Washington DC, 1984).
40. Roedder, E. *Am. Miner.* **69**, 413–439 (1984).
41. Holser, W. T. in *Symposium on Salt (First)* (ed. Berstiker, A. C.) 86–95 (Northern Ohio Geological Society, Cleveland, 1963).
42. McCaffrey, M. A. thesis, Univ. Harvard (1985).
43. Starinsky, A. thesis, Hebrew Univ. of Jerusalem (1974).
44. Zherbitsova, I. K. & Volkova, N. N. *Geochim. Int.* **3**, 656–670 (1966).
45. Stoffers, P. & Kühn, R. *Init. Rep. DSDP* **23**, 821–847 (1974).
46. Lazar, B. & Holland, H. D. (in preparation).
47. Lazar, B., Friedmann, T. J. & Holland, H. D. *Abstr. Prog. geol. Soc. Am.* **17**, 640 (1985).
48. Nielsen, H. *Geol. Rdsch.* **55**, 160–172 (1965).
49. Holser, W. T. & Kaplan, I. R. *Chem. Geol.* **1**, 93–135 (1966).
50. Claypool, G. E., Holser, W. T., Kaplan, I. R., Sakai, H. & Zak, I. *Chem. Geol.* **28**, 199–260 (1980).
51. Veizer, J., Holser, W. T. & Wilgus, C. K. *Geochim. cosmochim. Acta* **44**, 579–587 (1980).
52. Scholle, P. A. & Arthur, M. A. *Bull. Am. Ass. Petrol. Geol.* **64**, 67–87 (1980).
53. Schidlowski, M., Eichmann, R. & Junge, C. E. *Precamb. Res.* **2**, 1–69 (1975).
54. Keith, M. L. & Weber, J. N. *Geochim. cosmochim. Acta* **28**, 1787–1816 (1964).
55. Peterman, Z. E., Hedge, C. E. & Tourtelot, H. A. *Geochim. cosmochim. Acta* **34**, 105–120 (1970).
56. Veizer, J. & Compston, W. *Geochim. cosmochim. Acta* **38**, 1461–1484 (1974).
57. Brass, G. W. *Geochim. cosmochim. Acta* **40**, 721–730 (1976).
58. Faure, G., Assereto, R. & Tremblay, E. L. *Sedimentology* **25**, 523–543 (1978).
59. Starinsky, A., Bielski, M., Lazar, B., Wakshal, E. & Steinitz, G. *Earth planet. Sci. Lett.* **47**, 75–80 (1980).
60. Burke, W. H. *et al. Geology* **10**, 516–519 (1982).
61. DePaolo, D. J. & Ingram, B. L. *Science* **227**, 938–941 (1985).
62. Holser, W. T. in *Patterns of Change in Earth Evolution* (eds Holland, H. D. & Trendall, A. F.) 123–144 (Springer, New York, 1984).
63. Holland, H. D. *Geochim. cosmochim. Acta* **37**, 2605–2616 (1973).
64. Garrels, R. M. & Perry, E. A. *Jr. in The Sea Vol. 5* (ed. Goldberg, E. D.) Ch. 9 (Wiley, New York, 1974).
65. Garrels, R. M. & Lerman, A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4652–4656 (1981).
66. Garrels, R. M. & Lerman, A. *Am. J. Sci.* **284**, 989–1007 (1984).
67. Von Damm, K. L., Grant, B. & Edmond, J. M. in *Hydrothermal Processes at Seafloor Spreading Centers* (eds Rona, P. A., Boström, K., Laubier, L. & Smith, K. L.) 369–389 (Plenum, New York, 1983).
68. Albarède, F., Michard, A., Minister, J. F. & Michard, G. *Earth planet. Sci. Lett.* **55**, 229–236 (1981).
69. Shaw, H. F. & Wasserburg, G. J. *Geochim. cosmochim. Acta* **49**, 503–518 (1985).
70. Keto, L. S. *Abstr. Prog. geol. Soc. Am.* **17**, 628 (1985).
71. Quinby-Hunt, M. S. & Turekian, K. K. *EOS* **64**, 130–132 (1983).
72. Bruland, K. W. in *Chemical Oceanography Vol. 8* (eds Riley, J. P. & Skirrow, G.) (Academic, London, 1983).
73. Vine, D. D. & Tourtelot, E. B. *Econ. Geol.* **65**, 253–272 (1970).
74. Holland, H. D. *Econ. Geol.* **74**, 1676–1680 (1979).
75. Peltola, E. *Bull. de la Comm. géol. Finl.* (1960).
76. Leventhal, J. S. & Hosterman, J. W. *Chem. Geol.* **37**, 239–264 (1982).
77. Tankard, A. J. *et al. in Crustal Evolution of Southern Africa* Ch. 4 (Springer, New York, 1982).
78. Ramdohr, P. *Trans. geol. Soc. S. Afr.* **61** (Annexure) (1958).
79. Liebenberg, W. R. *Neues Jb. Miner. Mh.* **94**, 831–867 (1960).
80. Schidlowski, M. *Prof. Pap. U.S. geol. Surv.* 1161–N (1981).
81. Hiemstra, S. A. *Trans. geol. Soc. S. Afr.* **71**, 1–65 (1968).
82. Utter, T. J. *sedim. Petrol.* **50**, 71–76 (1980).
83. Grandstaff, D. E. *Precamb. Res.* **13**, 1–26 (1980).
84. Derry, D. R. *Econ. Geol.* **55**, 906–927 (1960).
85. Nicolaysen, L. O., Burger, A. J. & Liebenberg, W. R. *Geochim. cosmochim. Acta* **26**, 15–23 (1962).
86. Pienaar, P. J. *Bull. geol. Surv. Can.* **83** (1963).
87. Roscoe, S. M. *Geol. Surv. Can. Pap.* **68-40** (1957).
88. Rundle, C. C. & Snelling, N. J. *Phil. Trans. R. Soc. A286*, 567–583 (1977).
89. Simpson, P. R. & Bowles, J. F. W. *Phil. Trans. R. Soc. A286*, 527–548 (1977).
90. Grandstaff, D. E. *Econ. Geol.* **71**, 1493–1506 (1976).
91. Walker, J. C. G. *et al. in Earth's Biosphere* (ed. Schopf, J. W.) Ch. 11 (Princeton University Press, 1983).
92. Berkner, L. V. & Marshall, L. C. *J. atmos. Sci.* **22**, 225–261 (1965).
93. Cloud, P. *Scient. Am.* **249**, 176–189 (1983).
94. Hart, M. H. *Icarus* **33**, 23–39 (1978).
95. Kasting, J. F. & Walker, J. C. G. *J. geophys. Res.* **86**, 1147–1148 (1981).
96. Levine, J. S., Augustsson, T. R. & Natarajan, M. *Origins Life* **12**, 245–259 (1982).
97. Levine, J. S. in *The Photochemistry of Atmospheres* (ed. Levine, J. S.) Ch. 1 (Academic, Orlando, 1985).
98. Walker, J. C. G. in *Evolution of the Atmosphere* (Wiley, New York, 1977).
99. Towe, K. M. *Nature* **274**, 657–661 (1978).
100. Schidlowski, M., Eichmann, R. & Junge, C. E. *Precamb. Res.* **2**, 1–69 (1975).
101. Dimroth, E. & Kimberley, M. M. *Can. J. Earth Sci.* **13**, 1161–1185 (1976).
102. Dimroth, E. & Lichtblau, A. P. *Neues Jb. Miner. Abh.* **133**, 1–22 (1978).
103. Broecker, W. S. *J. geophys. Res.* **75**, 3553–3557 (1970).
104. Holland, H. D. & Zbinden, E. A. in *Physical and Chemical Weathering in Geochemical Cycles* (eds Lerman, A. & Meybeck, M.) (Reidel, Dordrecht, 1986).
105. Eckhardt, F. E. W. in *The Chemistry of Weathering* (ed. Drever, J. I.) (Reidel, Dordrecht, 1985).
106. Zika, R., Saltzman, E., Chameides, W. L. & Davis, D. D. *J. geophys. Res.* **87**, 5015–5017 (1982).
107. Kasting, J. F., Holland, H. D. & Pinto, J. P. *J. geophys. Res.* **90**, 497–510 (1985).
108. Buyanovsky, G. A. & Wagner, H. G. *J. Soil Sci. Soc. Am.* **47**, 1139–1145 (1983).
109. Reardon, E. J., Allison, G. B. & Fritz, P. J. *J. Hydrol.* **43**, 355–371 (1979).
110. de Jong, E. & Schappert, H. J. V. *Soil Sci.* **113**, 328–333 (1972).
111. Campbell, J. A. & Frascarelli, L. *Can. J. Soil Sci.* **61**, 137–144 (1981).
112. Rightmire, C. T. *Wat. Resources Res.* **14**, 691–692 (1978).
113. Richter, J. & Grossbauer, A. *Z. Pflanzenernaehr. Bodenkd.* **141**, 181–202 (1978).
114. Brook, G. A., Folkoff, M. E. & Box, E. O. *Earth Surf. Proc. Landforms* **8**, 79–88 (1983).
115. Gonchar-Zaykin, P. P. & Shvytov, I. A. *Pochovovdenie* (transl.), 65–72 (1973).
116. Albertsen, M. Z. *Pflanzenernaehr. Bodenkd.* **142**, 39–56 (1979).
117. Owen, T., Cess, R. D. & Ramanathan, V. *Nature* **277**, 640–642 (1979).
118. Hirst, D. M. *Am. Ass. Petrol. Geologists Mem.* **20**, 430–455 (1974).
119. Volkov, I. I. & Fomina, L. S. *Am. Ass. Petrol. Geologists Mem.* **20**, 456–476 (1974).

Diamond genesis in a multiply-constrained model

Stephen E. Haggerty

Department of Geology, University of Massachusetts, Amherst, Massachusetts 01003, USA

Thermal cracking of primordial carbon gas species in depleted and reduced Archaean subcratonic lithosphere is proposed for the nucleation of micro- and ultramafic macrodiamonds. Growth was mainly by solid-state diffusion, although crystallization of the eclogitic diamond suite was from sulphide-immiscible melts. Lateral and vertical heterogeneities and the geometry of the lithosphere can account for a variety of second-order variables such as diamond type, colour, crystal morphology and dissolution characteristics.

STUDIES of diamond genesis have been mostly concerned with the temperature (T) and pressure (P) environments of growth in the subcontinental lithosphere¹; comparatively little attention has been paid to the role of oxygen fugacity (f_{O_2}), which determines whether carbon exists as elemental carbon or as a carbon gas species (for example, CO, CO₂, CH₄). Basaltic and komatiitic volcanism and crustal extraction from the fertile garnet lherzolite asthenosphere has left an overlying depleted lithosphere of harzburgite and dunite which is estimated to be approximately 150–200 km thick in stable cratonic regions¹. Depletion included the extraction of volatiles, specifically H₂O + CO₂, with a consequent lowering of oxidation state². Contrasts in reduction–oxidation (redox) conditions between the lithosphere and the asthenosphere permit qualitative estimates to be made for the probability of diamond nucleation, and for subsequent dissolution and reaction. Diamonds are dated at 2,000–4,000 Myr^{3–6}, but sampling by kimberlites or lamproites occurred much later, between 1,200 and 20 Myr^{7–9}. The inferred antiquity of the lithosphere, as well as its heterogeneity and the interaction of oxidized asthenosphere with relatively reduced lithosphere, can account for the extreme variation in second-order characteristics of diamonds from a single diatreme, pipe cluster or diamond province.

Pressure–temperature–redox

Geothermometry of coexisting mineral inclusions in diamond¹ yields a temperature range of 900–1,400 °C. Pressures are estimated at 50–60 kbar (150–200 km). The carbon phase diagram¹⁰ includes a metastable region of predominantly diamond (+ graphite), and a field of predominantly graphite (+diamond), so that diamond may be grown metastably in the graphite stability field. Graphite is also a common inclusion in diamond¹¹.

Lithospheric redox conditions are in the range between the equilibria defined by the iron–wüstite (IW) and wüstite–magnetite (WM) buffers. The asthenosphere is more highly oxidized; redox conditions here lie between WM and FMQ (fayalite–magnetite–quartz)². Diamond and graphite stabilities, as a function of P – T – f_{O_2} in the COH system for 0.1 mol H₂O (ref. 12), are summarized in Fig. 1A. The carbon gas species are CO₂ and CH₄ at, respectively, high and low values of f_{O_2} . Diamond and graphite are in equilibrium with the system COH over a range of intermediate f_{O_2} values; diamond is stable at lower values than is graphite, and the equilibria change as functions of f_{O_2} , P and T . These data are displayed in Fig. 1B along with the FMQ, WM and IW buffers for a profile through the lithosphere and asthenosphere. Elemental carbon stability is embraced by the hachured lines, with diamond the dominant species at $T > 1,140$ °C and $P > 45$ kbar, and graphite at lower values. For asthenospheric conditions of WM–FMQ, diamond is unstable (closely spaced horizontal lines in Fig. 1B). However, if conditions are more reduced than WM, or the lithosphere–asthenosphere boundary is shallower than ~180 km, or the geothermal gradient is perturbed, a narrow range in f_{O_2} may be

present in which asthenospheric diamonds will crystallize. The situation is markedly different in the lithosphere, where the diamond stability region expands at the expense of the fields of CH₄ and CO₂, and where the redox state of WM–IW is entirely conducive to diamond crystallization. The conclusion that diamonds have grown in a reduced lithosphere is corroborated by diamond inclusions of metallic iron¹³, sulphides^{14,15}, magnesiowüstite¹⁶, Cr²⁺-bearing olivine^{17,18} (and possibly spinel¹⁹) and methane²⁰, with low to absent contents of Fe³⁺.

Diamond genesis

Figure 2 shows a schematic representation of the subcontinental lithosphere for a typical cratonic region containing diamond-bearing and non-diamond-bearing kimberlites and lamproites. Each craton and associated lithosphere differs in detail, but, as illustrated (Fig. 2), the lithosphere cross-section is modelled as keel-shaped. A basic assumption is that the lithosphere developed in the early Archaean and that the craton and associated lithosphere were tectonically stabilized by ~2,000 Myr^{21–23}. Another assumption is that crustal accretion developed very rapidly²⁴ and mainly by vertical underplating²⁵. Depleted lithosphere extended to depths that reached the diamond stability field by ~3,000 Myr^{1,6}. Whether Archaean Wilson-cycle tectonics of opening and closing basins or lateral ensialic accretion is favoured^{21–23}, the presence of mobile belts^{23,25} and sutures²⁶ surrounding anchored cratonic nuclei is unequivocal. Marginal ocean basins or sutures are high-heat-flow regions and imply lithospheric thinning. Present-day passive continental margins are modelled as having thin lithospheres at the boundary²⁷. By analogy, we suggest the existence of lithospheric thinning at craton–mobile belt boundaries²⁸. The slope of the lithosphere–asthenosphere boundary may be steeper than illustrated in Fig. 2, as is the case for the southern margin of the West African craton²⁹. However, from teleseismic^{30,31}, heat flow^{32,33} and electrical conductivity studies³⁴, the mobile belts are comparable to those in other parts of Africa^{35–38} and north-east South America³⁹. Gently dipping lithospheric contacts with the asthenosphere are also possible, but thicknesses of 150–200 km must be attained within the boundaries of the cratonic nucleus to support the growth of diamond. Therefore, a keel-shape, perhaps modified by convective sculpturing, is favoured. Concave isotherms and the form of the diamond–graphite stability curve are constrained by lithospheric shape. The base of the lithosphere is assumed to be isothermal, at ~1,300 °C (ref. 40). The lithosphere has a lower density, a lower f_{O_2} , is cooler, depleted and more rigid than the surrounding asthenosphere.

The anchored lithosphere must be a relatively effective barrier to asthenosphere-generated partial melts, with continuous vertical access being possible only in extraordinary tectonic circumstances. Because the base of the lithosphere is a region of contrasting thermal and f_{O_2} regimes, it is an ideal location for an imbalance in the equilibria between upwardly migrating gas species. As noted in Fig. 1, diamond nucleation requires oxi-

duction of CH_4 or reduction of CO_2 ; both may readily occur at the lithosphere–asthenosphere boundary. Gaseous cracking at this interface is proposed, with the precipitation of carbon and vapour fractionation of O–H. The mechanism is equivalent to the precipitation of carbon in the cooler segment of gas-mixing experimental furnace muffle tubes, and analogous to the soot that develops on the base of a cool receptacle over an open flame. In principle it is similar to diamond synthesis from a plasma of $\text{H}_2 + \text{CH}_4$ (refs 41, 42). It is this high-pressure soot, incorporated into and underplated onto the lithosphere, intergranular and finely dispersed, which is a possible source of carbon for diamonds. Carbon is a silicate-incompatible element and a second source of carbon is the dissociation of CO_2 or CH_4 during partial melting, with elemental carbon relegated to the lithospheric restite. Metamorphic growth may have continued for several million years, but barriers to intergranular diffusion may have caused some carbon to remain in the ultrafine particulate state as observed in studies of upper mantle xenoliths⁴³. Mantle degassing must have been particularly active in the interval before diatreme eruption, and the large concentration of particulate carbon that may have condensed is an ideal point source for microdiamonds. Transient eruptive temperatures were high, the growth and residence periods short, and the spontaneous nucleation of sharply faceted, corrosion-free, high-temperature octahedral diamonds and skeletal framesites is expected. These late-stage vapour condensates are, therefore, equivalent in age to, or only slightly older than, the eruptive event. Microdiamonds of >3,000 Myr age may exist, but only if securely protected from corrosive action and diffusional growth. Macrodiamonds of ultramafic affinity in the lithosphere are 2,000–4,000 Myr old, are metamorphic and grew in a volatile and geochemically depleted environment.

The genesis of eclogites is enigmatic and highly controversial⁷ but is receiving increased attention because eclogites and olivine-bearing eclogites (piclogites) are shown to satisfy seismic velocity contrasts in geophysical models of the upper and lower mantle⁴⁴. Eclogite compositions are, on average, equivalent to basalts but cannot be produced by anhydrous partial melting of garnet lherzolite at very high pressures⁴⁵; melt compositions⁴⁶, with increasing pressure, are basaltic (8–10 wt% MgO at ~10–15 kbar), picritic (~15–20 wt% MgO at 20 kbar) or komatiitic (>25 wt% MgO at >45 kbar). Diamond-bearing eclogites, therefore, cannot have been produced by simple anhydrous partial melting of fertile asthenosphere. A possible explanation for the origin of eclogites may be sought in the growth of the very earliest Archaean crust and lithosphere. Low-pressure melts trapped in the developing lithosphere or underplated onto the lower crust would have been largely basaltic in composition. Metamorphic transformation of basalt (gabbro) to eclogite results in an increase in density and salinity temperature relative to garnet lherzolite⁴⁵. Gravitational sinking⁴⁷ of eclogite (~3.5 g cm⁻³) may be more rapid through depleted lithosphere (~3.3 g cm⁻³) than in the underlying fertile asthenosphere (~3.4 g cm⁻³), leading to the accumulation of pods of eclogite at the lithosphere–asthenosphere boundary (Fig. 2). Depending on the depth, graphite, diamond or the metastable coexistence of diamond and graphite⁴⁸ might be expected. The ages of some eclogitic diamonds may thus approach the age of eclogite emplacement.

Although sulphur is a volatile gas, the evidence from xenoliths shows that sulphides are maintained in the upper mantle; they are also abundant as inclusions in diamonds^{14,15}. Sulphur may be an important catalyst in the upper mantle, reacting with hydrogen (to form H_2S) and carbon (to form COS or CS_2) to moderate the CO_2 : H_2 ratio and, hence, f_{O_2} (ref. 49). A solid–gas reaction of the form $2\text{FeS} + \text{CO}_2 = 2\text{FeO} + \text{S}_2 + \text{C}$ has been suggested as a mechanism for diamond nucleation⁵⁰. Spherical aggregates of equilibrated sulphides (pyrrhotite, pentlandite and chalcopyrite) are common in eclogites⁵¹, and have the appearance of immiscible liquids which could arise during the gravita-

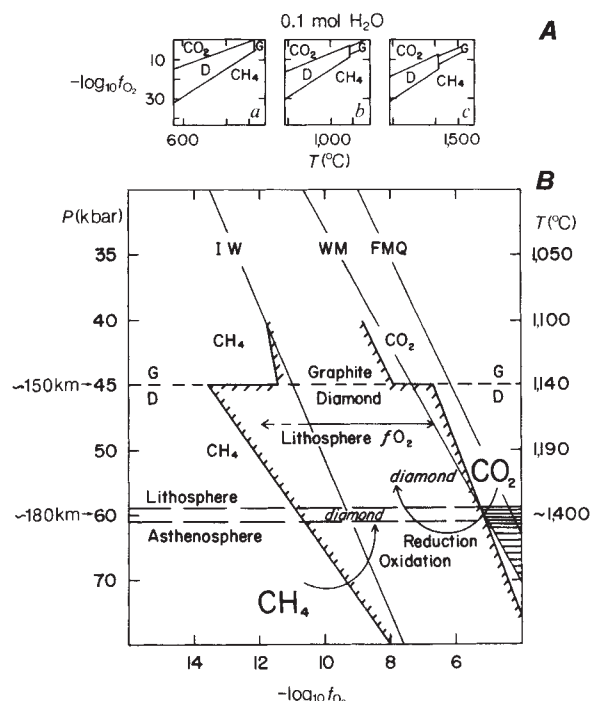
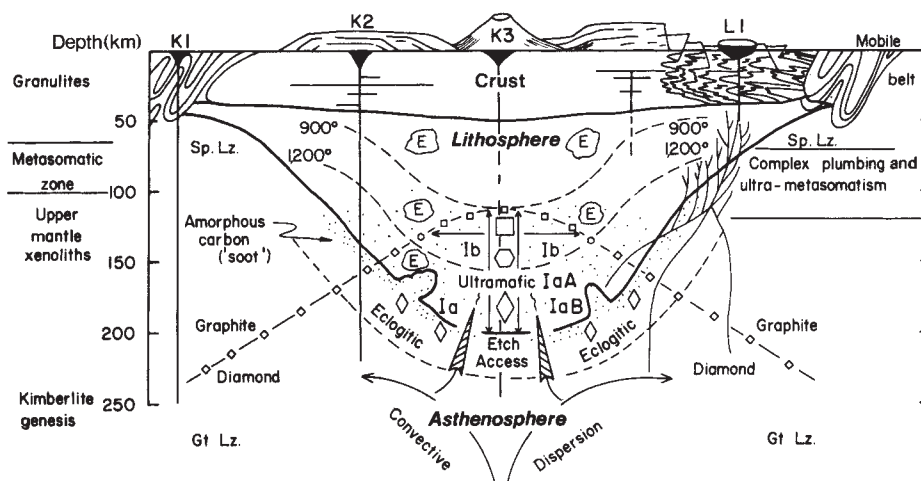


Fig. 1 A, Diamond–graphite relations in the system COH as a function of P , T and f_{O_2} in the presence of 0.1 mol H_2O (ref. 12). This is considered an appropriate model system for volatiles in the upper mantle. a, $P = 60$ kbar; b, $P = 50$ kbar; c, $P = 40$ kbar. CO_2 is stable at high values of f_{O_2} , CH_4 at lower values, and the elemental species over an intermediate range of f_{O_2} . In B, these data are integrated with the standard oxygen buffers, iron–wüstite (IW), wüstite–magnetite (WM) and fayalite–magnetite–quartz (FMQ), and the diamond–graphite P – T curve, in a cross-section of the lithosphere–asthenosphere boundary. A shield-type geothermal gradient is assumed. The hatched region (transcribed from a, b and c above¹²) is the zone in the upper mantle in which either diamond or graphite is stable relative to CH_4 or CO_2 . Redox estimates for the lithosphere², which are supported by diamond inclusions^{13–19} and gas-release studies²⁰, range from WM to some value <IW (double-headed arrow). This range is entirely in the diamond stable region at depths in the lithosphere >150 km. In the more highly oxidized asthenosphere (WM–FMQ), CO_2 is stable relative to diamond (horizontal shading). For diamond nucleation from distinct gas species, oxidation of CH_4 or reduction of CO_2 is required. A transition zone at ~180 km is shown at the lithosphere–asthenosphere boundary in which diamonds of asthenospheric affinity may crystallize.

tional sinking of eclogite to depths at which temperatures exceed ~1,100 °C, the liquidus of Fe–Ni–Cu monosulphide solid solutions⁵². Solubility of carbon would be greatly enhanced in these liquids and either diamond or graphite would crystallize on saturation.

From this viewpoint, eclogitic diamonds are melt-derived, with growth continuing by solid-state diffusion. Although not all diamonds contain sulphides and not all carbon was produced by sulphur or sulphide interaction, large concentrations of sulphides may have been retained in the upper mantle during subcontinental lithosphere development, aiding in the maintenance of a reduced buffered environment, critical to diamond nucleation and growth. The eclogitic diamond suite was perhaps most dependent on sulphide liquid immiscibility, and ultramafic-inclusion diamonds on sulphur reduction of CO_2 . This difference could account for variations in carbon and sulphur isotopes^{15,53–55}, and has a bearing on Ni:Fe ratios that vary with paragenesis, being higher in the ultramafic suite and correspondingly lower in eclogitic sulphides⁵⁶.

Fig. 2 A schematic representation of a typical subcratonic lithosphere with a crustal thickness of ~40 km and bounded by mobile belts. Depths at which melts are generated and xenolith horizons are listed to the left of the diagram. Rock types in the asthenosphere are spinel lherzolite (Sp.Lz.) and garnet lherzolite (Gt.Lz.). Note that the former is restricted to the margins of the lithosphere at shallow asthenosphere depths. The asthenosphere has a higher oxygen fugacity (f_{O_2}), density (ρ) and temperature (T), and is geochemically fertile and rheologically ductile relative to the depleted lithosphere, which is dominated by harzburgite and subordinate dunite. Concave isotherms are shown in the lithosphere at 900 and 1,200 °C, and the diamond-graphite curve is, therefore, convex. Symbols on the diamond-graphite curve and in the lithosphere depict high- T octahedral crystal forms ($\diamond$), intermediate- T cubo-octahedrons ($\circ$) and lower- T cubes ($\square$). It is argued (see Fig. 1 and text) that diamond nucleation is unlikely in the asthenosphere because of elevated f_{O_2} , but if diamonds do form they are likely to be recycled by convective dispersion. 'Ultramafic' in the lithosphere refers to the ultramafic inclusion suite in diamonds (Cr-pyroxene, olivine, enstatite, Cr-spinel and rare Cr-diopside). Irregular eclogite pods ('E') are shown sinking through the lithosphere and accumulating at the lithosphere-asthenosphere contact. The large hatched etch-access arrows at the base of the lithosphere are fracture paths for infiltrating partial melts and oxidized asthenosphere, considered to be responsible for diamond etching, partial modification or complete dissolution. Speckling represents the distribution of high-pressure vapour-deposited carbon (soot), microdiamonds or graphite in and on the cooler lithosphere. Diamond types (Roman numeral I) are nitrogen-bearing and are either annealed (a), with nitrogen in A and B clusters at low- and high- T , respectively, or are unannealed (b) in the cooler interior of the lithosphere. Vents K1, K2, K3 are typical kimberlite sampling profiles. L1 is considered probable for lamproites in north-west Australia, with a complex plumbing system, advanced metasomatism and intrusion into a thrust-sheeted mobile belt.



Nitrogen has been observed in all gas release studies of crushed diamonds^{57,58}, and spectral absorption type Ia diamonds dominate⁵⁹. Although generally regarded as an inert gas species, nitrogen may play an important role in diamond crystallization as outlined by Simakov⁶⁰, who has shown thermodynamically that methane interaction with nitrogen ($3CH_4 + 2N_2 = 3C + 4NH_3$) is energetically feasible for diamond precipitation. He has also proposed that high- T octahedra crystallize under more reducing conditions than do lower- T cubic morphologies. The N_2 - CH_4 reaction is attractive, considering that type I diamonds are so common, and that primordial degassing produced large concentrations of N_2 in the Earth's atmosphere. In common with S, it is proposed that N_2 may have acted catalytically to enhance carbon precipitation.

Variations in P and T are inherent in the geometry of the lithosphere and in its relation to the asthenosphere (Fig. 2). Isotherms are affected by thermal contrasts in the two units. If T varies, f_{O_2} varies, and the diamond stability field is modified. Basal and contact zones of the lithosphere will have remained at elevated temperatures and annealed high- T octahedral diamonds of type Ia are expected to dominate in these zones. We would expect cubo-octahedrons to develop towards the interior of the lithosphere, and restricted type Ib diamonds in the uppermost interior portions of the lithosphere. B-aggregated nitrogen centres⁵⁹ will form near the lithosphere base and A-clustered centres⁵⁹ should be prevalent in the cooler portions. Type II diamonds are restricted to isolated regions enriched in boron and depleted in nitrogen. Growth morphologies, diamond type and annealing characteristics should be similar for ultramafic and eclogitic suites.

The lithosphere is the residual product of asthenospheric partial melting and melt extraction, and heterogeneities in the abundances of trace elements (including carbon) in the restite are inevitable. As basalt and komatiite compositions vary, so must the lithospheric residue; hence, the observed variations in diamond chromophores, diamond size and lithospheric diamond grades. Heterogeneities will cause mechanical strength variations, with diamond deformation in the lower lithosphere

and trigon etching in regions of fractured lithosphere. A spectrum of resorbed morphologies (rounded dodecahedral or tetrahedral crystal forms) is expected, along with serial growth. It is generally assumed that macrodiamond dissolution results from interaction with transporting kimberlite⁶¹. This is unlikely because the event is transient and it is inconsistent with the ubiquitous presence in kimberlite of microdiamonds, which would be totally resorbed given their high surface-to-volume ratio. As noted above, vapour deposition of microdiamonds is favoured, closely related to the eruptive event.

Regional implications

Four contrasting diatreme intrusion paths are shown in Fig. 2, representing kimberlites (K1-3) and a lamproite (L). Intrusion K1, in the mobile belt adjacent to the craton, will sample fertile garnet and spinel lherzolites, and lower crustal lithologies. Intrusion K2 will sample a trapped eclogite enclave hosting diamonds, but the xenolith suite will include fertile and depleted ultramafic rocks and eclogites. Vent K3 will yield a typical lithospheric macrodiamond suite, and garnet lherzolites, harzburgites and dunites. In this model, the xenolith inclusion suite is an exploration indicator for the presence or absence of diamonds.

Lamproite intrusions of north-west Australia^{8,9} are located in the mobile belts surrounding the Kimberley craton. The diamond suite has eclogitic and ultramafic affinities^{9,62}, the lamproites have high contents of alkali and silicate-incompatible elements, and xenoliths have undergone a transition from garnet to spinel lherzolite. A complex plumbing system is envisaged, accompanied by advanced metasomatism. Long residence times in the metasomatic region of the lithosphere can account for the presence of dominantly small, highly modified and poor-quality diamonds. Spectral absorption data for diamonds from the Argyle lamproite show that most diamonds are in the IaB-aggregate form, but a complex annealing history is inferred, making these diamonds unlike any other diamonds previously studied⁶².

Lithospheric development is as essential to macrodiamond

formation as relatively impervious rigidity is to diamond survival. The lithosphere is a thermal and mechanical barrier, but tensile integrity may be compromised through gaseous dissociation and interaction, thus aiding in the liberation of diamonds from host rocks and accounting for the rarity of diamond-bearing xenoliths. Chemical interaction, furthermore, is evident from the disparity in compositions between encapsulated diamond inclusions and inferred host rocks^{6,7}. A developing lithosphere must be isostatically compensated and vertical fluctuation into and out of the diamond stability field over geological time is an inevitable consequence of the addition of surface volcanic piles or removal of crust by erosion. The region most sensitive to modification is at the diamond-graphite boundary. Here, diamond-minimum and graphite-maximum P-T conditions prevail, implying that the metastable coexistence of two carbon polymorphs may be more common than generally recognized⁴⁹. Diamonds growing in this low-*T* region have a minimal chance of survival, hence the relative rarity of low-*T* cubes as macrodiamonds and low-*T* unannealed type Ib diamonds.

The fact that Mesozoic intrusions in Africa contain macrodiamonds of Archaean age may be used to throw some light on the apparent absence or rarity of diamonds in many South American kimberlites. In Gondwanaland, lithospheres developed on the African and South American plates from a common, well-mixed asthenosphere. Precambrian kimberlite eruptions gave rise to the diamonds of Premier and Diamantina, and possibly others in southern, central and west Africa, and in north-east South America. Repositories of diamonds had already formed at >3,000 Myr. Continental fragmentation was initiated at ~200 Myr, drift continued and kimberlites were emplaced at 90–120 Myr. Africa remained essentially stationary but the lithospheric keel of South America was propelled through turbulently convective, oxidized asthenosphere. Torque on the lower lithosphere resulted in delamination⁶³, extensive brittle fracture, disruption, asthenospheric melt-fluid-gas infiltration and destruction of the diamond inventory. This scenario may be applicable to North America, India, Antarctica and Australia. In the case of north-west Australia, diamonds appear to have survived in a ductile and intensely metasomatized lithosphere.

Conclusions

First-order constraints of *P* and *T* alone are insufficient to define the conditions of nucleation and growth of diamonds in the upper mantle in the presence of COH system gas species. Diamond stability is controlled additionally by f_{O_2} and only strongly reducing environments meet the desired range of f_{O_2} conditions at elevated temperatures. Relatively reduced conditions stabilize carbon as CH₄; more oxidizing conditions stabilize CO₂. Redox conditions in the range WM to IW stabilize diamond in the lithosphere, relative to gaseous phases. Sulphur, either as COS or CS₂ or as sulphide-immiscible liquids, and N₂, are considered important as catalysts in carbon reduction, thus explaining the abundance of sulphide inclusions in diamond, and the widespread occurrence of type I diamonds. Other characteristics, such as growth and solution morphologies, colour, diamond type, carbon isotope abundances, size, serial growth, and deformation are explained by lateral and vertical variations in the lithosphere and by the recognition that diamonds grow by different processes. Lithospheric ultramafic diamonds are metamorphic and nitrogen-aggregated. Eclogitic diamonds nucleated in sulphur-rich melts and growth continued under metamorphic conditions, yielding some of the largest type II stones known. Macrodiamonds are older than 3,000 Myr and exhibit multiple growth, variable dissolution histories and the dominance of type Ia annealed stones. Growth continued over several million years by solid-state diffusion. Microdiamonds, other than those securely trapped and protected, are considered to be closely related in age to the eruptive event and are predicted

to be either type Ib or type IIa due to spontaneous crystallization and growth, and rapid quenching. Minimal dissolution takes place in the eruptive event, although clearly if conditions are excessively oxidizing, entrained diamonds will burn. Equally, if the f_{O_2} of the lithosphere is modified, for example, by infiltration of oxidized (WM-FMQ) asthenosphere, diamonds will be severely corroded and may not survive. Development of a sub-continental lithosphere appears to be essential to diamond formation, creating a reduced environment, acting as a barrier for eclogite entrapment, and providing a thermal discontinuity for gas disproportionation. Carbon is derived from primordial degassing and vapour fractionation of CH₄ and CO₂ in the presence of H₂O, S and N₂, by cracking within, or by underplating on to the cooler lithosphere. High-pressure carbon and diamond vapour condensates (soot) are invoked for the origin of microdiamonds, generated shortly before the eruptive event. The presence of macrodiamonds in exploration may be determined from depleted lithosphere xenolith suites and from high Mg+Cr signatures in satellite minerals of garnet, spinel and ilmenite.

This research was supported by the NSF (EAR83-08297) and by visiting fellowships from CSIR (University of Cape Town) and CSIRO (Mineral Chemistry Division, Melbourne) which are gratefully acknowledged. We thank A. J. Erlank, J. J. Gurney, H. O. A. Meyer, N. V. Sobolev, P. J. Wyllie, L. A. Bursill, L. A. Tompkins, D. Svisero, P. B. Toft and D. Leonard for useful discussions and F. R. Boyd, J. W. Harris, D. N. Robinson and C. B. Smith for reviews. We do not agree on every point but it is hoped that a framework is established for future research.

Received 24 June; accepted 19 December 1985.

- Boyd, F. R., Gurney, J. J. & Richardson, S. H. *Nature* **315**, 387–389 (1985).
- Haggerty, S. E. & Tompkins, L. A. *Nature* **303**, 295–300 (1983).
- Melton, C. E. & Giardini, A. A. *Geophys. Res. Lett.* **7**, 461–464 (1980).
- Ozima, M., Zashu, S. & Nitoh, O. *Geochim. cosmochim. Acta* **47**, 2217–2224 (1983).
- Kramers, J. D. *Earth planet. Sci. Lett.* **42**, 58–70 (1979).
- Richardson, S. H., Gurney, J. J., Erlank, A. J. & Harris, J. W. *Nature* **310**, 198–202 (1984).
- Dawson, J. B. *Kimberlites and Their Xenoliths* (Springer, Berlin, 1980).
- Atkinson, W. J., Hughes, F. E. & Smith, C. B. in *Kimberlites and Related Rocks: Developments in Petrology* Vol. 9 (ed. Kornprobst, J.) 195–225 (Elsevier, Amsterdam, 1984).
- Hall, A. E. & Smith, C. B. in *Kimberlite Occurrence and Origin: A Basis for Conceptual Models in Exploration* (eds Glover, J. E. & Harris, P. G.) 167–212 (University Extension, University of Western Australia, 1984).
- Bundy, F. P. *J. chem. Phys.* **38**, 631–635 (1963).
- Harris, J. W. & Gurney, J. J. in *The Properties of Diamond* (ed. Field, J. E.) 556–591 (Academic, New York, 1979).
- Deines, P. *Geochim. cosmochim. Acta* **44**, 943–961 (1980).
- Sobolev, N. V., Efimova, E. S. & Pospelova, L. M. *Geol. geophys. J. S.S.S.R.* **12**, 25–29 (1981).
- Harris, J. W. & Gurney, J. J. *Terra Cognita* **2**, 201 (1982).
- Tsai, H.-M., Shieh, Y. & Meyer, H. O. A. in *The Mantle Sample: Inclusions in Kimberlites and Other Volcanics* Vol. 2 (eds Boyd, F. R. & Meyer, H. O. A.) 87–103 (AGU, Washington, DC, 1979).
- Scott-Smith, B. H., Danchin, R. V., Harris, J. W. & Stracke, K. J. in *Kimberlites and Related Rocks: Developments in Petrology* Vol. 9 (ed. Kornprobst, J.) 121–142 (Elsevier, Amsterdam, 1984).
- Burns, R. G. *Contr. Miner. Petrol.* **51**, 213–221 (1975).
- Meyer, H. O. A. *Geochim. cosmochim. Acta* **39**, 929–936 (1975).
- Haggerty, S. E. in *The Mantle Sample: Inclusions in Kimberlites and Other Volcanics* Vol. 2 (eds Boyd, F. R. & Meyer, H. O. A.) 185–196 (AGU, Washington, DC, 1979).
- Giardini, A. A. & Melton, C. E. *Forsch. Miner.* **52**, 455–464 (1975).
- Anhaeusser, C. R. A. *Rev. Earth planet. Sci.* **3**, 31–53 (1975).
- Kröner, A. *Rev. Earth planet. Sci.* **13**, 49–74 (1985).
- Baer, A. J. *Tectonophysics* **72**, 203–227 (1981).
- Reymer, A. & Schubert, G. *Tectonics* **3**, 63–78 (1984).
- Kröner, A. in *Structure and Evolution of the Continental Lithosphere* Vol. 15 (eds Pollack, H. N. & Murthy, V. R.) 69–106 (Pergamon, New York, 1984).
- de Beer, J. H., Gough, D. I. & van Zijl, J. S. V. *Nature* **255**, 678–670 (1975).
- Sclater, J. G. & Francheteau, J. *Geophys. J. R. astr. Soc.* **20**, 509–542 (1970).
- Nisbet, E. G. *Can. J. Earth Sci.* **21**, 1426–1441 (1984).
- Haggerty, S. E. *J. geophys. Res.* **87**, 10, 811–10, 826 (1982).
- Briden, J. G. *et al. Nature* **292**, 123–128 (1981).
- Briden, J. G., Mereu, R. F. & Whitcombe, D. N. *Geophys. J. R. astr. Soc.* **71**, 793–808 (1982).
- Chapman, D. S. & Pollack, H. N. *Nature* **250**, 477–478 (1974).
- Brigaud, F., Lucazeau, F., Ly, S. & Sauvage, J. F. *Geophys. Res. Lett.* **12**, 549–552 (1985).
- Ritz, M. J. *geophys. Res.* **89**, 11, 317–11, 331 (1984).
- Fairhead, J. D. & Reeves, C. V. *Earth planet. Sci. Lett.* **36**, 63–76 (1977).
- van Zijl, J. S. V. *Trans. geol. Soc. S. Afr.* **81**, 129–142 (1978).
- Gough, D. I. in *Evolution of the Earth* (eds O'Connell, R. J. & Fyfe, W. S.) *Geodynamics Series* Vol. 5, 87–95 (AGU, Washington, DC, 1981).
- de Beer, J. H., van Zijl, J. S. V. & Gough, D. I. *Tectonophysics* **83**, 205–225 (1982).
- Wernick, E. *Earth Sci. Rev.* **17**, 31–48 (1981).
- Davies, G. F. *Earth planet. Sci. Lett.* **44**, 231–238 (1979).
- Matsumoto, S., Sato, Y., Tsutsumi, M. & Setaka, N. *J. Mater. Sci.* **17**, 3106–3112 (1982).
- Kamo, M., Sato, Y., Matsumoto, S. & Setaka, N. *J. Cryst. Growth* **62**, 642–644 (1983).
- Mathez, E. A. & Delaney, J. R. *Earth planet. Sci. Lett.* **56**, 217–232 (1981).
- Anderson, D. L. *Science* **223**, 347–355 (1984).

45. Yoder, H. S. Jr *Generation of Basaltic Magma* (NAS, Washington, DC, 1976).
46. Takahashi, E. & Scarfe, C. M. *Nature* **315**, 566–568 (1985).
47. Ito, K. in *High Pressure Research: Applications in Geophysics* (eds Manghni, M. H. & Akimoto, S.) 129–150 (Academic, New York, 1977).
48. Robinson, D. N. in *The Mantle Sample: Inclusions in Kimberlites and Other Volcanics* Vol. 2 (eds Boyd, F. R. & Meyer, H. O. A.) 50–58 (AGU, Washington, DC, 1979).
49. Haggerty, S. E. & Tompkins, L. A. in *Kimberlites and Related Rocks: Developments in Petrology* Vol. 9 (ed. Kornprobst, J.) 335–357 (Elsevier, Amsterdam, 1984).
50. Marx, P. C. *Miner. Mag.* **38**, 636–638 (1972).
51. Desborough, G. A. & Czamanske, G. K. *Am. Miner.* **58**, 195–202 (1973).
52. Craig, J. R. & Kullerud, G. *Econ. Geol. Monogr.* **4**, 344–358 (1969).
53. Deines, P., Gurney, J. J. & Harris, J. W. *Geochim. cosmochim. Acta* **48**, 325–342 (1984).

54. Sobolev, N. V., Galimov, E. M., Ivanovskaya, N. N. & Yefimova, E. S. *Dokl. Akad. Nauk SSSR* **249**, 1217–1220 (1979).
55. Milledge, H. J. *et al. Nature* **303**, 791–792 (1983).
56. Efimova, E. S., Sobolev, N. V. & Pospelova, L. N. *Zap. vses. miner. Obshch.* **117**, 300–310 (1983).
57. Melton, C. E. & Giardini, A. A. *Am. Miner.* **60**, 413–417 (1975).
58. Melton, C. E. & Giardini, A. A. *Am. Miner.* **66**, 746–750 (1981).
59. Bibby, B. M. *Chem. & Phys. Carbon* **18**, 1–91 (1982).
60. Simakov, S. K. *Dokl. Akad. Nauk SSSR* **268**, 182–186 (1983).
61. Robinson, D. N. *Miner. Sci. Engng* **10**, 55–72 (1978).
62. Harris, J. W. & Collins, A. T. *Ind. Diam. Rev.* **3**, 128–130 (1985).
63. McKenzie, D. & O'Nions, R. K. *Nature* **301**, 229–231.

LETTERS TO NATURE

Does dark matter affect the solar neutrino flux?

A. De Rújula*, S. L. Glashow† & L. Hall†

* Theory Division, CERN, CH-1211 Geneva 23, Switzerland

† Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

An observed deficit of high-energy (^8B) solar neutrinos¹ and the galactic dark-mass problem² have both been explained in terms of hypothetical particles (ref. 3 and L. M. Krauss, unpublished work): 'cosmions' must have masses m_c in the range 5–50 AMU, a mean elastic scattering cross-section per hydrogen nucleus in the Sun of $\langle\sigma\rangle_s \sim 4 \times 10^{-36} \text{ cm}^2$, local mass density $\rho \sim 1 \text{ AMU cm}^{-3}$, and cosmological velocity $v \sim 10^{-3}c$. Accumulated in the Sun, they enhance its thermal conductivity so as to yield the measured solar luminosity with a lower central temperature than in standard models. Here we show that such particles, should they exist, could be detected and studied in the laboratory by bolometric procedures proposed originally by Cabrera *et al.*⁴ for the purpose of measuring solar neutrinos. The event rates are so large that a definitive test of this cosmion hypothesis can be made, even with a relatively small bolometric detector.

According to Cabrera *et al.*⁴ neutrinos from p-p reactions in the Sun produce recoil electrons in Si at the rate of $10^{-3} \text{ kg}^{-1} \text{ day}^{-1}$. Higher-energy ^8B neutrinos scatter coherently from Si nuclei depositing several keV of energy at a rate one-seventh as large. Cabrera *et al.*⁴ have proposed an experiment using tons of silicon to study these effects. Cosmions with properties specified in ref. 3 also make elastic collisions with Si nuclei depositing up to 1.3 keV (for $m_c = 5 \text{ AMU}$) and up to 23 keV (for $m_c = 50 \text{ AMU}$). In general, the maximum energy transfer is $2Mm_c v^2 / (M + m_c)^2$ where M is the mass of the target nucleus. Cosmion scattering is S-wave so that the distribution in recoil energy is uniform.

The cosmion cross-section on a nucleus is coherent as its de Broglie wavelength exceeds the nuclear size. For simplicity, we take cosmion-nucleus scattering to be isospin independent, and comment later on the effects of relaxing this constraint.

For a cosmion with purely spin-dependent (SD) interactions, the non-relativistic scattering cross-section on element i is related to that on hydrogen by⁵:

$$\sigma_i(\text{SD}) = \left\langle \left[\frac{m_c + 1}{m_c + A} \right]^2 A^2 \frac{4}{3} J(J+1) \right\rangle_i \sigma_H(\text{SD})$$

where m_c is in AMU, A and J are the mass number and nuclear spin of an isotope of the element and the average is taken over the various isotopes according to their natural abundance. In the Sun, such cosmions scatter predominantly from hydrogen, so that cosmions can provide the link between the dark mass of the galactic halo and the solar (^8B) neutrino (ν) flux providing $\sigma_H(\text{SD}) \approx 4 \times 10^{-36} \text{ cm}^2$. This leads to an event rate per kg per

day in a bolometric detector of element i of

$$\sim 8,300 \left\langle \frac{A}{m_c} \left[\frac{m_c + 1}{m_c + A} \right]^2 J(J+1) \right\rangle_i$$

For this case, Si is a poorly chosen target; only 4.7% of naturally occurring Si has a nuclear spin, and these are only spin 1/2, giving an event rate of 53(70) per kg per day for $m_c = 5(50) \text{ AMU}$. For a 5 AMU cosmion, there are several elements which give event rates 200 times larger than Si: Li, B, Al, Sc, V and Nb, while for a heavy cosmion (50 AMU) these elements give event rates relative to Si which increase with atomic number ranging from 50 for Li to 800 for Nb.

For spin independent (SI) cosmion-nucleus scattering⁵

$$\sigma_i(\text{SI}) \approx \left[\frac{m_c + 1}{m_c + A} \right]^2 A^4 \sigma_H(\text{SI})$$

where A is the dominant mass number for element i . The large growth of $\sigma_i(\text{SI})$ with A means that for all values of m_c of interest, cosmion scattering in the Sun is dominated by He and the heavy elements rather than by H. To obtain $\langle\sigma_s\rangle = 4 \times 10^{-36} \text{ cm}^2$, we find $\sigma_H(\text{SI}) \sim 2 \times 10^{-37} \text{ cm}^2$ for $m_c = 5 \text{ AMU}$ and $\sim 2 \times 10^{-38} \text{ cm}^2$ for $m_c = 50 \text{ AMU}$. Event rates per kg per day are

$$\sim 1.6 \times 10^{39} \sigma_H(\text{SI}) \frac{A^3}{m_c} \left[\frac{m_c + 1}{m_c + A} \right]^2$$

which for Si is 4.6×10^4 (6×10^3) for $m_c = 5(50) \text{ AMU}$. A further enhancement by a factor of 10 in rate is obtained with a target element of $A \sim 200$.

In particle physics models of the cosmion, it is likely that the scattering from nuclei will have some isospin dependence. This will change our numerical estimates but is unlikely to affect our conclusions.

Event rates for bolometric detection of various sources and targets is shown in Table 1. Entries relating to solar neutrinos or reactor antineutrinos near the core are taken from ref. 4. They are to be compared with our estimates of cosmion-induced events. These event rates are so high that they will be observable

Table 1 Event rates in a bolometric detector

	(kg ⁻¹ day ⁻¹)	
Recoil electrons from solar p-p ν s on Si	1×10^{-3}	
Coherent scattering of solar ^8B ν s on Si nuclei	1.5×10^{-4}	
$\bar{\nu}$ events on Si near a nuclear reactor core	40	
	$m_c = 5 \text{ GeV}$	$m_c = 50 \text{ GeV}$
Cosmion scattering on Si nuclei (spin dependent)	53	70
Cosmion scattering on B nuclei (spin dependent)	1.4×10^4	1.7×10^4
Cosmion scattering on B nuclei (spin independent)	1.2×10^4	6.0×10^2
Cosmion scattering on Si nuclei (spin independent)	4.6×10^4	6.0×10^3

with a small detector providing sensitivity to deposition energies of 1–20 keV is obtained. A significant signal is still expected when probing incident cosmions with velocities near the galactic escape velocity of $2 \times 10^{-3} c$, in which case the recoil energies will be 4–80 keV. These event rates also apply to superconducting colloid detectors⁶ which could also detect cosmions.

Cosmions present a far easier target for bolometric detection than do solar neutrinos. The Press–Spergel hypothesis can be tested decisively. Should the existence of cosmions be demonstrated, bolometric measurements (in Si, B and possibly other materials) can yield their mass, local density, velocity distribution and the magnitude and spin-dependence of cosmion–nucleon cross-sections.

We thank members of the Physics Department of Boston University where this research was done.

Received 31 October; accepted 20 December 1985.

1. Davis, R. in *Proc. Telemark Neutrino Mass Miniconf.* (eds Barger, V. & Cline, D. University of Wisconsin Press, Madison, 1980).
2. Rubin, V. C. *Science* **220**, 1339–1344 (1983).
3. Press, W. H. & Spergel, D. N. *Astrophys. J.* **296**, 679–684 (1985).
4. Cabrera, B., Krauss, L. M. & Wilczek, F. *Phys. Rev. Lett.* **55**, 25 (1985).
5. Goodman, M. & Witten, E. Preprint, Princeton University (1985).
6. Drukier, A. & Stodolsky, L. *Phys. Rev. D* **30**, 2295 (1984).

The real value of Mercury's perihelion advance

Anna M. Nobili*† & Clifford M. Will‡

* School of Mathematical Sciences, Theoretical Astronomy Unit, Queen Mary College, London E1 4NS, UK

‡ McDonnell Center for the Space Sciences, Department of Physics, Washington University, St Louis, Missouri 63130, USA

The perihelion advance of the orbit of Mercury has long been one of the observational cornerstones of general relativity. After the effects of the gravitational perturbations of the other planets have been accounted for, the remaining advance fits accurately to predictions according to the theory of general relativity, that is, 43 arc s per 100 yr. In fact, radar determinations of Mercury's motion since the mid-1960s have produced agreement with general relativity at the level of 0.5%, or to ± 0.2 arc s per 100 yr. We have recently been puzzled by an apparent uncertainty in the value of the theoretical prediction for the advance within general relativity, as cited in various sources. The origin of this discrepancy is a 1947 article by Clemence¹, who used unconventional values for the astronomical unit (A) and the speed of light (c) in his calculation of the predicted advance. Since that time, virtually everyone has followed suit. Although the current value of the prediction, using the best accepted values for the astronomical constants and for the orbital elements of Mercury, is 42.98 arc s per 100 yr, there is a preponderance of citations over the past 25 years of Clemence's value 43.03 arc s per 100 yr. Here we derive the accurate value and uncover the source of Clemence's value.

According to general relativity, the predicted perihelion precession rate ($\dot{\omega}$) is given by²

$$\dot{\omega} = \frac{6\pi GM_{\odot}}{Pa(1-e^2)c^2} \quad (1)$$

where G is the newtonian gravitational constant, $M_{\odot}$ is the mass of the Sun and a , e and P are the semi-major axis, eccentricity and orbital period of Mercury, respectively, whose mass has been neglected with respect to $M_{\odot}$. However, to obtain a numerical value for $\dot{\omega}$, note that $M_{\odot}$ cannot be measured directly. Instead the product $GM_{\odot}$ is determined through measurements

of the orbital periods (or mean motions) and the semi-major axes of planetary orbits, and the use of Kepler's third law. In relativistic language, $GM_{\odot}$ is said to be the 'Kepler-measured' mass, as distinct from the total rest mass of all the particles contained in the Sun. This Kepler-measured mass is given by

$$GM_{\odot} = n^2 a^3 \quad (2)$$

where $n = 2\pi/P$ is the mean motion of a body of negligible mass in an orbit of semi-major axis a . It is conventional to define the gaussian gravitational constant k , given by

$$k^2 \equiv \bar{n}^2 \bar{a}^3 \quad (3)$$

where $\bar{n}$ is the mean motion in radians per mean solar day of a body of negligible mass with respect to $M_{\odot}$ (unit mass) and $\bar{a}$ is its semi-major axis in AU. By international agreement, k has been given the defining value of 0.01720209895 since 1938. This is the value first assigned to it by Gauss on the basis of values available at that time for the Earth's orbital period and for the Earth–Sun mass ratio. This value serves to define A as the radius of a circular orbit around the Sun on which a body of negligible mass would complete one revolution in $2\pi/k = 365.2568983$ mean solar days. The solar mass is then related to A by

$$GM_{\odot} = k^2 A^3 / D^2 \quad (4)$$

where A is in metres, D = mean solar day = 86,400 s and $GM_{\odot}$ is expressed in $\text{m}^3 \text{s}^{-2}$. The semi-major axes of planetary orbits are usually measured in units of A , so if we define $\bar{a} = a/A$, we obtain for $\dot{\omega}$

$$\dot{\omega} = \frac{6\pi k^2}{PD^2 \bar{a}(1-e^2)} \left[\frac{A^2}{c^2} \right] \quad (5)$$

Now, k and D are defined constants, while for Mercury, the mean orbital elements $\bar{a}$, e and P are known accurately and have maintained constant values to better than 1 part in 10^5 for at least the past 80 years. Those values are $\bar{a} = 0.387099$, $e = 0.205630$ and $P = 0.24085$ yr. From the International Astronomical Union (IAU) 1976 System³ (Table 1), the current adopted values for A and c are

$$A_{1976} = 1.4959787 \times 10^{11} \text{ m} \quad (6)$$

$$c_{1976} = 299,792,458 \text{ m s}^{-1} \quad (7)$$

In terms of these values, we obtain

$$\dot{\omega} = 42.98 (A/A_{1976})^2 (c/c_{1976})^{-2} \text{ arc s per 100 yr} \quad (8)$$

However, A and c have not always had their current values. Before 1964, the value of the astronomical unit was a *derived* quantity, obtained from the solar horizontal parallax (the angle subtended by the Earth's equatorial radius at 1 AU), which was treated as a primary constant, and had the adopted value of 8.8 arc s, and from the radius of the Earth, also treated as a primary constant with the value 6.378388×10^6 m (see Table 1). The derived value of A was 1.495042×10^{11} m. But by the mid-1960s, radar ranging to the inner planets produced direct and more accurate determinations of distance (through measurements of light travel time) and of A , and it was found that the previous value was too small by $\sim 90,000$ km. By 1964, IAU had adopted the radar value as a primary constant, and relegated the solar parallax to the status of a derived constant⁴. Minor refinements in the value of A were made in the 1976 IAU system. Similarly, the value of c that was used before the mid-1960s was $299,860,000 \text{ m s}^{-1}$, a value dating from Michelson's measurements of the late nineteenth century. It was replaced by the lower value $299,792,500 \text{ m s}^{-1}$, based on modern measurements.

With the older values of A and c , the predicted perihelion advance before 1964 would have been 42.91 arc s per 100 yr. With the current values, the prediction is 42.98 arc s per 100 yr. Where then did the value 43.03 come from?

† Permanent address: Dipartimento di Matematica, Università di Pisa, I-56100 Pisa, Italy.

This value originates from an article¹ by Clemence which reviews the observations of the overall perihelion shift, the subtraction of the general precession of the geocentric coordinate system and the removal of the contribution from planetary perturbations. However, Clemence's analyses adopted what were at the time unconventional values of A and c . Instead of the accepted value of 8.8 arc s for the solar parallax, he chose a value of 8.790 arc s, which, combined with the conventional Earth radius, gave a value for A of 1.4967429×10^{11} m. This value is slightly closer to the current value, although $\sim 0.05\%$ too high, than was the conventional 1947 value, which was about 0.06% too low. Furthermore, for c he also chose a non-conventional value which he took from a review by Dorsey⁵ of measurements of c . This value, $299,773,000 \text{ m s}^{-1}$, was lower than the Michelson value, although quite close to the current value. The other constants for Mercury's orbit were the standard values. His result for the general relativistic prediction was 43.03 arc s per 100 yr.

Clemence's value was generally accepted for 35 years. In many textbooks and review articles on general relativity, authors quote the rounded-off values 43 or 43.0 arc s per 100 yr, both of which are compatible with either Clemence's or the present-day value. However, of those authors who quote four-figure accuracy, most agree with Clemence. A survey of the literature shows 18 publications since 1960 containing Clemence's value, but only 4 with the modern value. The distribution of these citations is plotted as a function of year in Fig. 1.

Two cases in point are the classic textbooks on gravitation by Misner, Thorne and Wheeler² and by Weinberg⁶. Both texts list in their appendices values for A and c which are consistent with the 1964 IAU System, yet both fall into step behind the Clemence perihelion prediction (p. 1113 in ref. 2; p. 198 in ref. 6). Clemence's¹ value is also maintained in other major textbooks⁷⁻¹². We have been unable to determine why Clemence's value was so widely used.

Figure 1 also shows the value 42.95 arc s per 100 yr quoted by Richard¹³ and by Will¹⁴⁻¹⁶. This value is anomalous and appears to be the product of rounding off one of the ingredients of the prediction (such as the orbital period) too severely before inserting into the formula.

Although of theoretical interest, the difference between these quoted predictions for Mercury's perihelion advance has no observational consequences, for the following reason. In the current method of testing gravitational theory using the dynamics of Mercury, the actual value of the predicted advance is irrelevant. Instead, the equations of motion used to determine the ephemerides of Mercury and the inner planets automatically include relativistic post-newtonian terms whose time-averaged consequence is the perihelion shift. The size or effect of these terms in the equations of motion is modulated by a set of dimensionless parameters, 'parametrized post-newtonian (PPN) parameters', whose values depend on the theory of gravitation being used (they have specific fixed values in general relativity)¹⁵. These parameters, termed, for example, γ , β , α_1 , become part of a multi-parameter least-squares fit procedure in which radar data and other appropriate measurements are used to improve the parameters in the least-squares sense^{17,18}. Other parameters include the usual ones, for example, initial conditions for the

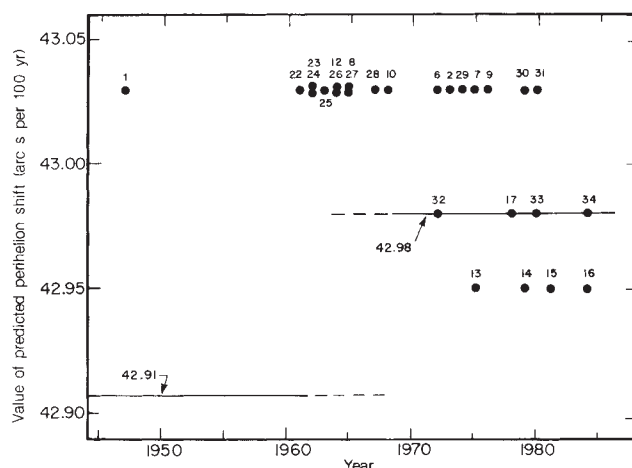


Fig. 1 Quoted values of the general relativistic prediction for Mercury's perihelion advance 1947-85. Horizontal lines are the values dictated by currently accepted values for A and c . Numbers above circles are references.

planets, station locations, A and the Earth-Moon mass ratio. The outcome of such experiments is an improved set of values for all the parameters. The values of the PPN parameters can then be compared with their predicted values in various theories. Thus, although it is possible that the accuracy of the determinations of the PPN parameters by this method may eventually reach an accuracy corresponding to the 0.05 arc s per 100 yr difference between the Clemence value for $\dot{\omega}$ and the true value, the actual value for the net perihelion precession never appears in the analysis. Current observations combining Mercury radar data with Viking Mars data are nearing this level of accuracy¹⁸, and a possible future Mercury orbiter may do even better¹⁹.

The interpretation of the perihelion advance is further complicated by the contribution of a newtonian quadrupole gravitational field generated by a solar oblateness. The possible contributions to $\dot{\omega}$ depend on the size of the solar mass quadrupole moment, as inferred from various observations or from theoretical models of the Sun. The contributions range in size from 3 arc s per 100 yr, inferred from the original Dicke-Goldenberg 1966 visual solar oblateness measurements, to 1 arc s per 100 yr, from more recent 1983 visual observations²⁰, to 0.6 arc s per 100 yr, inferred from some solar oscillation measurements²¹, to 0.1 arc s per 100 yr, inferred from Mercury/Viking Mars data¹⁸, to 0.01 arc s per 100 yr, inferred from standard, uniformly rotating theoretical solar models.

Nevertheless, the real general relativistic perihelion shift of Mercury can be stated unequivocally: it is 42.98 arc s per 100 yr.

This research was supported in part by NSF grant PHY83-13545, by the Italian Research Council and by NATO Research grant 737/84.

Received 6 December 1985; accepted 7 January 1986.

1. Clemence, G. M. *Rev. mod. Phys.* **19**, 361-364 (1947).
2. Misner, C. W., Thorne, K. S. & Wheeler, J. A. *Gravitation* (Freeman, San Francisco, 1973).
3. *The Astronomical Almanac for the year 1984* (US Government Printing Office, Washington, DC, 1983).
4. *Supplement to the Astronomical Almanac 1968* (US Government Printing Office, Washington, DC, 1966).

Table 1 Astronomical constants and Mercury's perihelion

Constant	1947 (ref. 35)	1947 (ref. 1)	1960 (ref. 35)	1964 IAU system ⁴	1976 IAU system ³
AU A (10^{11} m)	(1.495042)	(1.4967429)	(1.495042)	1.49600	1.4959787
Earth radius (m)	6,378,388	6,378,388	6,378,388	6,378,160	6,378,140
Parallax (arc s)	8.8	8.790	8.8	(8.794)	(8.794148)
c (m s^{-1})	299,860,000	299,773,000	299,860,000	299,792,500	299,792,458
$\dot{\omega}$ (arc s per 100 yr)	42.91	43.03	42.91	42.98	42.98

Values in parentheses are derived constants.

5. Dorsey, N. E. *Trans. Am. phil. Soc.* **34**, 1-106 (1944).
6. Weinberg, S. *Gravitation and Cosmology* (Wiley, New York, 1972).
7. Adler, R., Bazin, M. & Schiffer, M. *Introduction to General Relativity* 2nd edn (McGraw-Hill, New York, 1975).
8. McVittie, G. C. *General Relativity and Cosmology* (University of Illinois, Urbana, 1965).
9. Ohanian, H. C. *Gravitation and Spacetime* (Norton, New York, 1976).
10. Robertson, H. P. & Noonan, T. W. *Relativity and Cosmology* (Saunders, Philadelphia, 1968).
11. Stephani, H. *General Relativity* (Cambridge University Press, 1982).
12. Synge, J. L. *Relativity: The General Theory* (North-Holland, Amsterdam, 1964).
13. Richard, J.-P. in *General Relativity and Gravitation* (eds Shaviv, G. & Rosen, J.) 169-188 (Wiley, New York, 1975).
14. Will, C. M. in *General Relativity: An Einstein Centenary Survey* (eds Hawking, S. W. & Israel, W.) 24-89 (Cambridge University Press, 1979).
15. Will, C. M. *Theory and Experiment in Gravitational Physics* (Cambridge University Press, 1981).
16. Will, C. M. *Phys. Rep.* **113**, 345-422 (1984).
17. Anderson, J. D., Keesey, M. S. W., Lau, E. L., Standish, E. M. Jr & Newhall, X X *Acta astr.* **5**, 43-61 (1978).
18. Hellings, R. W. in *General Relativity and Gravitation* (eds Bertotti, B., de Felice, F. & Pascolini, A.) 365-385 (Reidel, Dordrecht, 1984).
19. Ashby, N., Bender, P. L. & Wahr, J. M. *Proc. 10th int. Conf. General Relativity and Gravitation* 947 (Conf. Pap., Padua, 1983).
20. Dicke, R. H., Kuhn, J. R. & Libbrecht, K. G. *Nature* **316**, 687-690 (1985).
21. Hill, H. A., Bos, R. J. & Goode, P. R. *Phys. Rev. Lett.* **49**, 1794-1797 (1982).
22. Tangherlini, F. R. *Nuovo Cim. Suppl.* **20**, 1-86 (1961).
23. Bertotti, B., Brill, D. R. & Krotkov, R. in *Gravitation: An Introduction to Current Research* (ed. Witten, L.) 1-48 (Wiley, New York, 1962).
24. Born, M. *Einstein's Theory of Relativity* (Dover, New York, 1962).
25. Pathria, R. K. *The Theory of Relativity* (Hindustan Publishing Corp., Delhi, 1963).
26. Dicke, R. H. *The Theoretical Significance of Experimental Relativity* (Gordon & Breach, New York, 1964).
27. Dicke, R. H. & Peebles, P. J. *Space Sci. Rev.* **4**, 419-460 (1965).
28. Dicke, R. H. *Phys. Today* **20**, No. 1, 55-70 (1967).
29. Rees, M., Ruffini, R. & Wheeler, J. A. *Black Holes, Gravitational Waves and Cosmology: An Introduction to Current Research* (Gordon & Breach, New York, 1974).
30. Ginzburg, V. L. *Sov. Phys. Usp.* **22**, 514-527 (1979).
31. Cowsik, R. in *Gravitation, Quanta and the Universe* (eds Prasanna, A. R., Narlikar, J. V. & Vishveshwara, C. V.) 18-40 (Wiley, New York, 1980).
32. Brumberg, V. A. *Relyativistskaya Nebesnaya Mekhanika* (Izdatelstvo Nauka, Moskva, 1972).
33. Will, C. M. *Ann. N.Y. Acad. Sci.* **336**, 307-321 (1980).
34. Straumann, N. *General Relativity and Relativistic Astrophysics* (Springer, Berlin, 1984).
35. *The American Ephemeris and Nautical Almanac* (US Government Printing Office, Washington, DC, 1947; 1960).

Extragalactic nature of G227.1+1.0

Gary A. Chanan*, David J. Helfand†, Hyron Spinrad‡ & Kate Ebner†

* Department of Physics, University of California, Irvine, California 92717, USA

† Columbia Astrophysics Laboratory, 538 W. 120th Street, New York, New York 10027, USA

‡ Astronomy Department, University of California, Berkeley, California 94720, USA

Green and Gull¹ have presented evidence that G227.1+1.0, one of the small-diameter radio sources in the Molonglo survey of Clark and Crawford², is a Crab-like supernova remnant. If this is the case, the object is an extremely important one, as Green and Gull emphasize¹. Such remnants are rare and are valuable as probes of the particle acceleration by pulsars. The very small diameter (~1') of G227.1+1.0 implies that it is one of the youngest of these remnants, perhaps only a few hundred years old. It could then represent the first known example of a Phase I Crab-like remnant³; that is, one for which the age is less than the slowing timescale of the underlying pulsar. Such a remnant would be uniquely important in helping to establish by observation an evolutionary picture of these objects (see ref. 4 for a theoretical discussion). Furthermore, the direction of G227.1+1.0—towards the galactic anti-centre—makes it a good candidate for optical study, as the usual problems of crowding and high extinction in the plane should be reduced. Here we report the discovery of a bright ($V = 17.4$), diffuse optical object within a few arc seconds of the centre of the radio source, but we present spectroscopic evidence that it is a luminous external galaxy at $z = 0.073$. We argue, furthermore, that the original radio emission represents not a supernova remnant, but a radio halo surrounding this galaxy, similar to that around M87.

Figure 1 is a V image of the G227.1+1.0 field taken with the prime-focus charge coupled detector (CCD) at the 4-m telescope of the Cerro Tololo Inter-American Observatory (CTIO) on 19 February 1985. The diffuse nature of the central object in the field is readily apparent. The centre of the radio emission (roughly 1' by 2' in extent) is not particularly well defined (there is no central point source), but lies roughly 0.1' west of the centre of the optical emission. The magnitudes of the diffuse object, uncorrected for reddening, were determined from this and other CCD exposures to be $U = 20.1$, $B = 18.9$, $V = 17.4$ and $I = 15.4$, all accurate to 0.1 mag.

Figure 2 shows a spectrum of the optical object obtained with the Lick 3-m CCD spectrograph on 20 March 1985. The spectrum is similar to that of M87 (refs 5, 6) and G127.1+0.5, except that the emission is weaker (here only [NII] at $\lambda 6,584$ is clearly seen in emission), which suggests that the object is a giant elliptical

galaxy. The redshift obtained from the emission line and the three absorption features labelled in the figure is $z = 0.073$. For such a galaxy, this redshift then implies $(B - V)_z = 1.2$ (refs 7, 8) so that in terms of magnitude, $V_0 = 16.5$ and $M_V = -21.7$ (for $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$), ~1.5 magnitudes fainter than M87. We note that the extinction along this line-of-sight through the galactic plane is remarkably low ($A_V \sim 0.9$) for $b = 1.0^\circ$.

The analogy to M87 is strengthened if we identify the radio emission reported by Green and Gull¹ with this external galaxy. M87 shows radio emission on a variety of angular scales^{9,10}, but the most relevant here is the halo, 12' by 16' in extent, with a mean surface brightness at 408 MHz of $3.2 \times 10^{-19} \text{ W m}^{-2} \text{ Hz}^{-1} \text{ sr}^{-1}$ (ref. 11). By comparison, the 1-GHz surface brightness of G227.1+1.0 reported by Green and Gull¹ extrapolates to almost the identical value at 408 MHz: $3.4 \times 10^{-19} \text{ W m}^{-2} \text{ Hz}^{-1} \text{ sr}^{-1}$, and this object is about a factor of 2 larger than M87 in spatial extent. This close analogy suggests strongly that the radio source G227.1+1.0 is associated with the optical galaxy, and does not represent a young galactic supernova remnant. The failure to detect a compact nuclear source in this object is perhaps not surprising in view of the

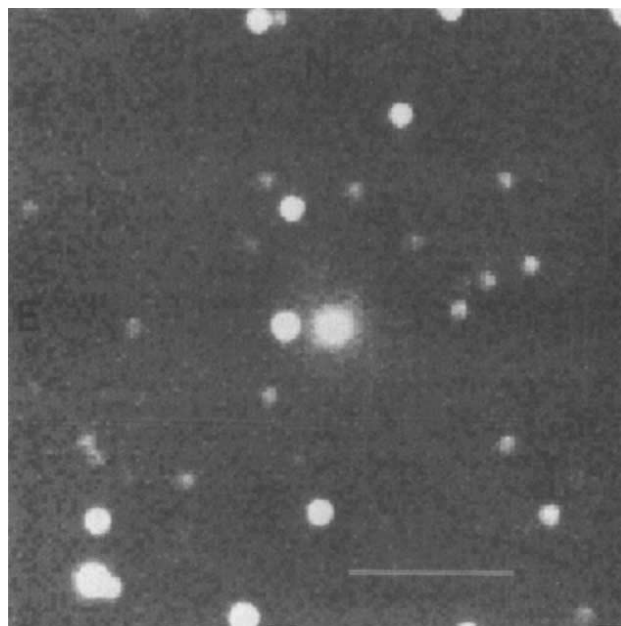


Fig. 1 V image of field of G227.1+1.0 taken with the prime-focus CCD at the 4-m telescope of CTIO. The extended nature of the central object, which we identify as a giant elliptical galaxy, is readily apparent. Scale bar represents 20" in length. Exposure time was 20 seconds.

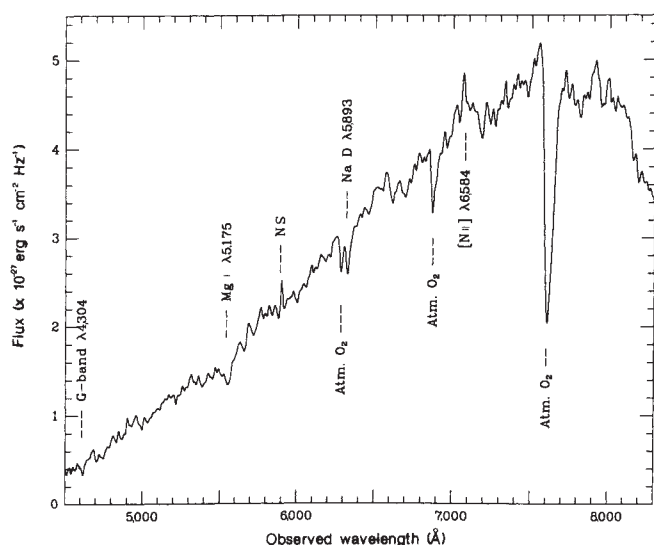


Fig. 2 Optical spectrum of the diffuse source coincident with G227.1+1.0 taken on 20 March 1985 UT with the CCD spectrograph at the Lick Observatory 3 m. Atm. O₂, atmospheric oxygen; NS, night sky.

relative weakness of the (presumably nuclear) optical emission. We note that the unresolved nuclear source in M87, if moved to $z=0.073$, would have an intensity of 14 mJy at 5 GHz. A further reduction of a factor of severalfold would probably render it difficult to detect against the background of the extended source.

That an external galaxy should appear in an in-plane survey is not surprising as the Molonglo survey on which the Green and Gull study was based covered a large area of the sky (0.40 sr) and, in fact, most of the original 513 objects in the survey are expected to be extragalactic². At the same time, the discovery of a bright optical galaxy within 6" of the centre of the radio source of interest is unlikely to be a chance superposition: there are about 30 galaxies per square degree brighter than $V=17.4$ (ref. 12), so that this latter probability is only $\sim 3 \times 10^{-4}$.

In retrospect there are several problems with the interpretation of G227.1+1.0 as a very young supernova remnant. The radio spectral index $\alpha=0.6$ reported by Green and Gull¹ is too large by a factor of ≥ 2 for a purely Crab-like component. It is conceivable that such a component could be suppressed by free-free absorption at ages ≤ 100 yr (ref. 4), leaving the steeper shell-like component to dominate, but this is not consistent with the structural interpretation of Green and Gull. On the other hand, such a large spectral index is typical of extragalactic objects; the total emission from M87 is described by $\alpha=0.8$. It is true, as Green and Gull note¹, that the observed surface brightness of $2 \times 10^{-19} \text{ W m}^{-2} \text{ Hz}^{-1} \text{ sr}^{-1}$ is typical of several of the young remnants (Crab-like and otherwise), but this is not a compelling argument: half of the ~ 100 radio galaxies tabulated by Lang¹³ have surface brightness at 1 GHz of this same order of magnitude. If one were to maintain the (in our view highly improbable) viewpoint that the optical galaxy is a chance superposition on a legitimate young radio supernova remnant, it follows that the supernova event which gave rise to the latter would have been readily visible at optical wavelengths despite its low galactic latitude, as we have shown that there is little material in the line-of-sight. By the same token, the lack of any evidence of an optical supernova remnant in the field seems significant.

Thus we conclude that the radio source G227.1+1.0 is not a supernova remnant, but an extragalactic object which we identify with the luminous and relatively unobscured galaxy dis-

cussed here. The observational history of this object therefore rather closely parallels that of G127.1+0.5 (see ref. 14 and refs therein). Most of the estimated 50 Crab-like remnants in the Galaxy with radio surface brightnesses that are predicted to be detectable⁴ thus continue to evade identification.

We thank B. Margon for useful discussions and for assistance with the data reduction, and M. Bolte for the astrometric measurements. G.A.C. and D.J.H. are Alfred P. Sloan Foundation Research Fellows, and G.A.C. is a Visiting Astronomer of the Cerro Tololo Inter-American Observatory, supported by the NSF under contract AST 78-27879. D.J.H. acknowledges the support of the U.S. Air Force (AFOSR 820014). This is UCI technical report 85-49.

Received 2 December 1985; accepted 8 January 1986.

1. Green, D. A. & Gull, S. F. *Nature* **312**, 527-529 (1984).
2. Clark, D. H. & Crawford, D. F. *Aust. J. Phys.* **27**, 713-727 (1984).
3. Pacini, F. & Salvati, M. *Astrophys. J.* **186**, 249-265 (1973).
4. Reynolds, S. P. & Chevalier, R. A. *Astrophys. J.* **630**, 648 (1984).
5. Stauffer, J. & Spinrad, H. *Astrophys. J. Lett.* **231**, L51-L55 (1979).
6. Ford, H. C. & Butcher, H. *Astrophys. J. Suppl.* **41**, 147-172 (1979).
7. Whitford, A. E. *Astrophys. J.* **169**, 215-220 (1971).
8. Schild, R. & Oke, J. B. *Astrophys. J.* **169**, 209-214 (1971).
9. Donaldson, W., Miley, G. K. & Palmer, H. P. *Mon. Not. R. astr. Soc.* **152**, 145-158 (1971).
10. Turland, B. D. *Mon. Not. R. astr. Soc.* **170**, 281-294 (1975).
11. Cameron, M. J. *Mon. Not. R. astr. Soc.* **152**, 439-460 (1971).
12. Allen, C. W. *Astrophysical Quantities* 3rd edn (Athlone, London, 1973).
13. Lang, K. R. *Astrophysical Formulae* 2nd edn (Springer, New York, 1980).

Neutral hydrogen absorption towards G227.1+1.0

D. A. Green & S. F. Gull

Mullard Radio Astronomy Observatory, University of Cambridge, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK

The radio source G227.1+1.0 (ref. 1) was observed in a search for young galactic supernova remnants, and its peculiar structure indicated that it might be a very young remnant. Chanan *et al.*², however, have recently found a bright optical object with $z=0.073$ close to the radio centre, which represents strong circumstantial evidence that the source is in fact a radio galaxy. Here we present a neutral hydrogen absorption profile observed with the Very Large Array (VLA) which confirms the extragalactic nature of this object.

To provide further information as to the nature of G227.1+1.0, and to measure its distance if galactic, we made neutral hydrogen observations towards this source using the VLA in 'D array' during October 1984. The resolution of the resulting maps was ~ 1 arc min, comparable in size to the object, which is appropriate for absorption studies. Thirty velocity channels covering the range -25 to 125 km s^{-1} (with respect to the local standard of rest) were available, together with continuum observations. Figure 1 shows the emission spectrum in the vicinity of G227.1+1.0 from the Berkeley survey (ref. 3), together with the absorption profile deduced from our VLA observations. Absorption can be seen against all spiral arms of our Galaxy.

Although Chanan *et al.*² have argued that the radio structure of G227.1+1.0 is similar to that of the halo of M87, we believe that it resembles more closely 3C310 (ref. 4) than any other radio source. However, these interpretations may be consistent if M87 and 3C310 are both examples of medium-power twin-jet radio sources in which the jet-axis lies relatively close ($\sim 20^\circ$) to the line of sight, and hence the large-scale lobes overlap in projection. In any case, there is now no doubt that G227.1+1.0 is a radio galaxy.

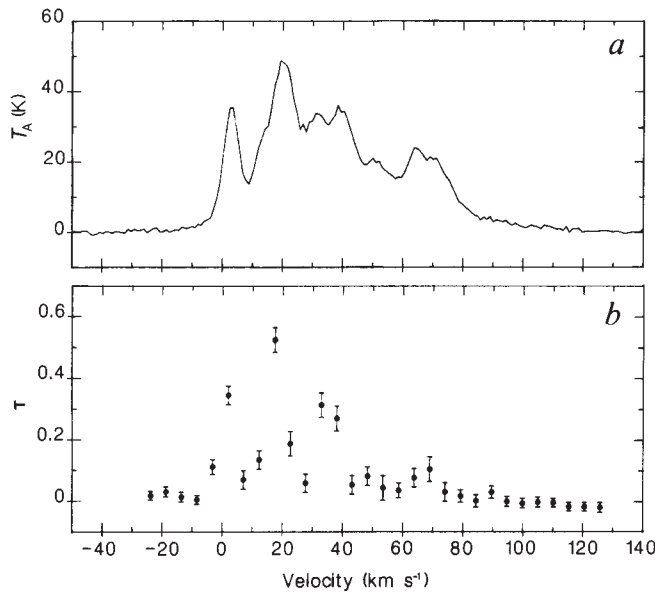


Fig. 1 *a*, Neutral hydrogen emission (T_A , antenna temperature) in the vicinity of G227.1+1.0, taken from Weaver and Williams³. *b*, Optical depth (τ) of absorption from our VLA observations. Velocity is relative to the local standard of rest.

The VLA is operated by Associated Universities Inc., under contract from the US NSF.

Received 30 December 1985; accepted 9 January 1986.

1. Green, D. A. & Gull, S. F. *Nature* **312**, 527-529 (1984).
2. Chanan, G. A., Helfand, D. J., Spinrad, H. & Ebner, K. *Nature* **320**, 41-42 (1986).
3. Weaver, H. & Williams, D. R. W. *Astr. Astrophys. Suppl.* **1**, 1-503 (1973).
4. van Breugel, W. J. M. & Fomalont, E. B. *Astrophys. J.* **282**, L55-L58 (1984).

High-radio-frequency survey for young and millisecond pulsars

T. R. Clifton* & A. G. Lyne

Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

A high-radio-frequency survey for distant pulsars is being carried out at Jodrell Bank. An area of 200 square degrees along the galactic plane has been observed with high sensitivity for objects having both short periods ($P \geq 0.004$ s) and large dispersion measures ($\leq 35,000$ P cm⁻³ pc). We have confirmed 32 new pulsars, including one of period 85 ms and two objects of longer period which both have characteristic ages of the order of 20,000 yr. Only three known pulsars are younger than this. Most of the new pulsars have very high dispersion measure (≥ 400 cm⁻³ pc), and many display significant multi-path scattering, which suggests that, even at this high radio frequency, we may still be missing a significant number of short-period pulsars which would otherwise be detected in a flux-limited sample. No new pulsars have been detected in binary systems or with 'millisecond' periods.

* Present address: Berkeley Astronomy Department, University of California, Berkeley, California 94720, USA.

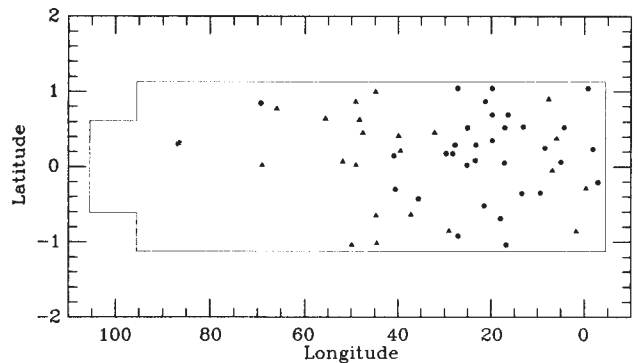


Fig. 1 The distribution in galactic coordinates of all pulsars detected in the Jodrell Bank B survey. Solid line, survey boundaries; circles, new pulsars; triangles, previously catalogued as pulsars.

The 400 known pulsars do not reflect the true population of pulsars in the Galaxy because of the observational biases associated with past surveys. In particular, those with short period or high dispersion measure (implying large distance) are under-represented. Undesirable selection effects have resulted from the relatively low sampling rates used and the dispersive broadening of pulsar signals within receiver channels, while multi-path interstellar scattering and high galactic background emission^{1,2} have also seriously limited previous large-scale surveys conducted at around 400 MHz.

Our experiment was performed to clarify the true distribution of short-period pulsars—particularly young objects such as the Crab and Vela pulsars—by overcoming those specific effects which would have prevented their detection in previous periodicity surveys. This is achieved in our new search, the Jodrell Bank B Survey, by sampling faster and using the much higher observing frequency of 1,400 MHz. The main disincentive for using such a high frequency is the small size of the telescope beam and the corresponding slow rate of survey. However, this choice of frequency does offer several important advantages. First, although pulsars have steep spectra with typical indices of ~ -2 at frequencies of a few hundred megahertz³, the galactic background radiation is even steeper, with a spectral index of ~ -2.6 (ref. 4). In those regions of the galactic plane where the receiver sensitivity is limited by galactic background emission, for example, towards the inner regions of the Galaxy, observations at high frequency will be more sensitive for most pulsars. Second, greater bandwidth is available at the higher frequencies, which also increases the sensitivity. Despite this increased bandwidth, dispersive broadening of the pulses is greatly reduced. Finally, at 1,400 MHz, the broadening of pulses caused by interstellar scattering is reduced by a factor of ~ 150 compared with observations at around 400 MHz.

Observations of the galactic plane in the region $-4^\circ \leq l \leq 105^\circ$, $|b| < 1^\circ$ were carried out at Jodrell Bank between April 1983 and October 1984 with the MKIA telescope at an observing frequency of 1,400 MHz. Two circularly polarized signals were received over a 40-MHz total bandwidth and independently filtered into 8 contiguous channels, each of width 5 MHz. The 16 detected outputs were then summed in polarization pairs to form 8 total power signals, each of which was integrated and sampled every 2 ms. After removal of the mean level, each data-stream was then scaled, digitized with three-bit accuracy and stored on magnetic tape. Cooled FET amplifiers provided a typical system noise of ~ 50 K, equivalent to 55 Jy with the MKIA radio telescope. With a dwell time of ~ 9 min on each point on the sky and allowing for various system losses, this provided a 6 σ detection level of ~ 1 mJy for a pulsar having a duty cycle of 4%.

We recorded data on 250 magnetic tapes which were processed on the Cyber 205 computer at the University of Manchester

Table 1 Parameters of 32 new pulsars

PSR	Period (JD 2,446,141) (ms)	Error	DM (cm ⁻³ pc)	Galactic coordinates		Error	$S_{1,400}$ (mJy)
				Long.	Lat.		
1736-31	529.437	1	620 ± 30	357.10	-0.21	0.05	5*†
1736-29	322.8806	4	139 ± 3	359.20	1.04	0.10	5
1737-30	606.5431	5	149 ± 3	358.28	0.23	0.03	6
1750-24	528.331	6	800 ± 60	4.30	0.52	0.10	4*†
1753-24	670.480	1	370 ± 20	5.00	0.06	0.05	2†
1800-21	133.57862	5	230 ± 15	8.40	0.25	0.03	13‡
1804-20	918.40	1	660 ± 70	9.40	-0.35	0.10	2*
1809-17	1,205.368	7	210 ± 20	13.11	0.53	0.05	2
1813-17	782.326	4	490 ± 30	13.40	-0.35	0.10	1†
1815-14	291.4888	8	595 ± 35	16.40	0.69	0.10	6*†
1817-13	921.43	2	750 ± 50	17.10	0.52	0.10	3*†
1819-14	214.7714	6	560 ± 100	17.20	-0.16	0.05	2
1820-11	279.822	6	430 ± 50	19.80	1.04	0.10	6
1821-11	435.758	5	620 ± 40	19.80	0.69	0.10	2*
1822-14	279.178	7	360 ± 40	16.80	-1.04	0.10	5
1823-11	2,093.09	5	460 ± 200	19.80	0.35	0.10	2
1823-13	101.4364	1	219 ± 10	18.00	-0.69	0.10	4§
1824-09	245.7576	4	410 ± 50	21.30	0.87	0.10	3*†
1829-08	647.275	2	286 ± 13	23.31	0.29	0.03	4
1829-10	330.354	4	440 ± 40	21.50	-0.52	0.10	2
1830-08	85.28106	1	402 ± 9	23.39	0.08	0.03	6*†
1832-06	305.83	1	410 ± 40	25.10	0.52	0.10	2*
1834-04	354.237	3	200 ± 30	27.20	1.04	0.10	2
1834-06	1,905.71	1	310 ± 80	25.18	0.02	0.05	5
1838-04	186.14511	9	304 ± 7	27.82	0.29	0.05	4¶
1839-04	1,839.95	2	200 ± 50	28.30	0.17	0.10	9
1841-05	255.6965	2	400 ± 12	27.07	-0.92	0.05	4*†
1842-02	507.718	3	410 ± 40	29.70	0.17	0.10	1
1855+02	415.8116	6	480 ± 30	35.60	-0.43	0.05	3
1903+07	648.039	1	380 ± 10	40.91	0.14	0.05	2
1904+06	267.2741	3	465 ± 10	40.61	-0.30	0.05	5#
2000+32	696.7314	3	135 ± 10	69.26	0.84	0.03	5

The error in galactic position applies to both coordinates, and for pulse periods, the error quoted applies to the last decimal place given. JD, Julian days; DM, dispersion measure.

* Measurable scattering at 1,400 MHz; † Measurable scattering at 920 MHz; ‡ $\dot{P} = 125 \pm 5 \times 10^{-15}$; § $\dot{P} = 76 \pm 6 \times 10^{-15}$; || $\dot{P} = 9 \pm 1 \times 10^{-15}$; ¶ $\dot{P} \leq 40 \times 10^{-15}$; # $\dot{P} \leq 90 \times 10^{-15}$.

Regional Computer Centre. The processing principally involved the execution of fast Fourier transforms on de-dispersed data for assumed dispersion measures in the range 0–17,000 cm⁻³ pc. A harmonic averaging algorithm similar to that used in the second Molonglo survey⁵ was used to select candidates from the spectra having periods of >4 ms. Signal-to-noise ratios were then determined (for all but the shortest periods) by folding the original de-dispersed data at the suspect periods. To reduce the large number of suspects per observation to a short-list suitable for re-observation, the Cyber 205 outputs were transferred to the VAX 11/780 computer at the Jodrell Bank STARLINK node and pulsar candidates selected on the basis of pulse profiles and signal-to-noise ratios. Re-observations were carried out at 1,400 MHz and most confirmed pulsars were subsequently also observed at 920 MHz.

We have now established 32 new pulsars (Table 1) and detected 22 previously known pulsars. We expect that when further telescope time becomes available ~10 more candidates will be confirmed as pulsars. None of the new pulsars shows evidence for any binary motion. Figure 1 indicates the positions of all pulsars detected in the search. Note the new highly dispersed pulsars below galactic longitude 40°, where previous surveys were insensitive because of the factors discussed above. Two of the stronger (and lower dispersion measure) objects, PSR1736-29 and PSR1737-30, were detected in extremely hot regions of the sky close to the direction of the galactic centre. This illustrates the ability of these observations to overcome the effects of sky

background emission, which severely restricted the sensitivity of previous lower-frequency observations at low galactic latitudes.

The absence of any pulsars at high longitude, despite maximum sensitivity in the region, clearly implies that pulsars are far less common at large galactocentric radius than in the inner regions of the Galaxy which have not really been penetrated before. When the survey is complete, our data will allow a more precise estimate of the enhancement of pulsar numbers within galactocentric radii of several kiloparsecs than has so far been possible.

The distribution in period of the new pulsars, compared with those previously catalogued (Fig. 2), is characterized by an increased proportion of faster objects ($P \leq 0.2$ s)—a result of the improved sensitivity in this region and the younger population expected near the galactic plane⁶. One or two longer-period pulsars ($P \geq 1$ s) may have been missed because of confusion with periodic interference from aviation radars.

The younger pulsars detected in the survey may be associated with detectable supernova remnants, and we have investigated this possibility for several objects. The pulsar of shortest period, PSR1830-08 ($P = 85$ ms), is not clearly associated with any known supernova remnant, although a possible association with W41 cannot be excluded because of the confused nature of the region. Period measurements made at epochs separated by 2 yr have constrained the period derivative ($\dot{P}$) to $(9 \pm 1) \times 10^{-15}$, assuming that the 'spot' measurements of period are not affected

by any binary motion. This suggests a characteristic age of $\sim 150,000$ yr, so that an association with a remnant such as those observed around the relatively young Crab and Vela pulsars is not expected.

$\dot{P}$ s have also been obtained (again, on the assumption of no binary motion) for the short-period objects PSR1823-13 ($P = 101$ ms) and PSR1800-21 ($P = 133$ ms), both yielding values of the order of 10^{-13} (Table 1). The implied characteristic ages are 22,000 and 17,000 yr, respectively, indicating that both are very young objects. Only three known pulsars are younger; Crab, Vela and PSR1509-52, all of which lie within known supernova remnants. These two new pulsars therefore seem the most promising candidates to have associated remnants. The dispersion measure of PSR1823-13 implies a distance of ~ 5.5 kpc, based on the galactic electron density distribution proposed by Lyne *et al.*⁶. This object does not appear to be closely associated with

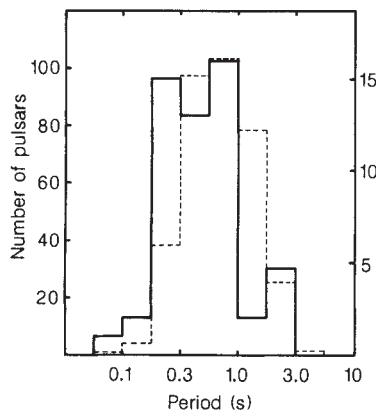


Fig. 2 A comparison of the distribution in period of those objects detected in the Jodrell Bank B survey (solid-line histogram and right-hand vertical scale) and those pulsars detected in previous major surveys (broken-line histogram and left-hand vertical scale).

any known remnant in the recent catalogue of Green⁷ and no radio morphology suggestive of a remnant is evident in the 2.7-GHz survey map of Reich *et al.*⁸. For PSR1800-21 the dispersion-derived distance is 5.3 kpc and again there is no obvious remnant association, although a region of emission of 50 arc min extent lies ~ 30 arc min to the west and a 6 arc min knot of emission is at a similar angular separation to the south. Note also that neither pulsar lies close to any of the X-ray sources of the Ariel V catalogue⁹, nor is there an association with any γ -ray source in the COS-B catalogue¹⁰.

Of the other newly detected pulsars, possible associations between PSR1838-04 and G27.8+06 (ref. 11) or PSR1904+06 and G40.0+0.5 (ref. 12) seem more promising from positional considerations. However, preliminary measurements (Table 1) indicate that neither pulsar has the high $\dot{P}$, and hence the small characteristic age, that would strengthen such an association.

The effects of interstellar scattering are evident in the new sample even at this relatively high radio frequency. Figure 3 presents integrated profiles of 10 of the new pulsars at both 1,400 and 920 MHz. The profiles of some pulsars (such as PSR1736-31) show the clear exponential tails and the strong frequency dependence characteristic of multi-path scattering. In Table 1 we have indicated those pulsars whose pulses are apparently broadened by interstellar scattering at 1,400 or 920 MHz. It is clear that many would be undetectable at significantly lower frequencies.

The pulsar PSR1736-29 shows both a main pulse and an interpulse (Fig. 3). The separation in longitude of the pulses is

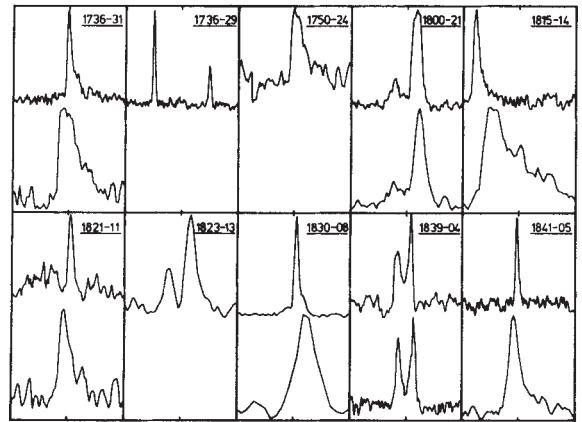


Fig. 3 Integrated pulse profiles for 10 of the more interesting pulsars discovered in the Jodrell Bank B survey. For each object, a pair of profiles is shown covering the whole pulse period: the upper obtained at a radio frequency of 1,400 MHz and the lower at 920 MHz where available. Only in the case of PSR1750-24 is the lack of a lower-frequency waveform indicative of a non-detection, presumably resulting from strong interstellar scattering.

very close to 180° , providing very strong evidence that, at least in this case, they are radiated from two opposite poles of the neutron star. Note also the asymmetrical, wide, two-component pulses of PSR1800-21, PSR1823-13 and PSR1839-04. In the first two, emission extends over approximately half of the pulse period; such profiles are unusual in pulsars of such short period.

In conclusion, our survey shows that the so-called millisecond pulsars are not common objects, in agreement with the recent results of Stokes *et al.*¹³. However, our search detected some short-period young objects which fit naturally into the evolutionary sequence of the main population of pulsars⁶. It is important to seek any supernova remnants that are associated with them. The observation of these pulsars makes those evolutionary models that involve the injection of long-period pulsars into the population less likely^{14,15}. Significant scattering, even at 1,400 MHz, is likely to inhibit the detection of pulsars with periods similar to that of the Crab pulsar in the inner regions of the Galaxy and beyond. The 400-MHz survey of Stokes *et al.*¹³ detected only one new pulsar within our survey boundaries. This illustrates the degrading effects of high sky background temperatures and interstellar scattering on searches of the galactic plane at low frequency, which have been partially overcome in the experiments described here.

We thank Dr M. Ashworth for his invaluable contribution to the development of the software on the CYBER 205 and Dr A. W. Jones for assistance with data processing. T.R.C. acknowledges financial support from SERC.

Received 18 November 1985; accepted 6 January 1986.

1. Lyne, A. G. in *Proc. Green Bank Workshop on Millisecond Pulsars* (eds Reynolds, S. P. & Stinebring, D. R.) 223-226 (National Radio Astronomy Observatory, 1984).
2. Manchester, R. N., D'Amico, N. & Tuohy, I. R. *Mon. Not. R. astr. Soc.* **212**, 975-986 (1985).
3. Manchester, R. N. & Taylor, J. H. *Pulsars* (Freeman, San Francisco, 1977).
4. Haslam, C. T. G., Salter, C. J., Stoffel, H. & Wilson, W. E. *Astr. Astrophys. Suppl.* **47**, 1-142 (1982).
5. Manchester, R. N. *et al. Mon. Not. R. astr. Soc.* **185**, 409-421 (1978).
6. Lyne, A. G., Manchester, R. N. & Taylor, J. H. *Mon. Not. R. astr. Soc.* **213**, 613-639 (1985).
7. Green, D. A. *Mon. Not. R. astr. Soc.* **209**, 449 (1984).
8. Reich, W., Furst, E. & Sofue, Y. *Astr. Astrophys.* **133**, L4-L7 (1984).
9. Warwick, R. S. *et al. Mon. Not. R. astr. Soc.* **197**, 865-891 (1981).
10. Swannenburg, B. N. *et al. Astrophys. J.* **243**, L69-L73 (1981).
11. Reich, W., Furst, E., Steffan, P., Reif, K. & Haslam, C. T. G. *Astr. Astrophys. Suppl.* **58**, 197-248 (1984).
12. Downes, A. J. B., Pauls, T. & Salter, C. J. *Astr. Astrophys.* **92**, 47-50 (1980).
13. Stokes, G. H., Taylor, J. H., Weisberg, J. M. & Dewey, R. J. *Nature* **317**, 787-788 (1985).
14. Vivekanand, M. & Narayan, R. J. *Astrophys. J.* **2**, 315-337 (1981).
15. Radhakrishnan, V. *Contemp. Phys.* **23**, 207-337 (1982).

Crystal structure determination by powder neutron diffraction at the spallation neutron source, ISIS

A. K. Cheetham*, W. I. F. David†, M. M. Eddy*,
R. J. B. Jakeman*, M. W. Johnson† & C. C. Torardi‡

* Chemical Crystallography Laboratory, University of Oxford,
9 Parks Road, Oxford, Oxon OX1 3PD, UK

† Rutherford-Appleton Laboratory, Chilton, Didcot,
Oxon OX11 0QX, UK

‡ E. I. Du Pont de Nemours and Co., Experimental Station,
Wilmington, Delaware 19898, USA

The power and scope of powder diffraction methods have advanced dramatically in recent years, partly as a consequence of improved X-ray and neutron instrumentation¹, and partly due to the development of Rietveld profile analysis² for the refinement of low-symmetry structures. Emphasis has focused on neutron studies, with both constant wavelength and time-of-flight techniques, but progress has also been made in X-ray methods³, where synchrotron sources promise to play an important part⁴. At present, structural models with more than 100 variable parameters can be refined^{5,6}, but the determination of structures from powder data has remained an empirical, rather than exact, science. The principal difficulty is that the overlapping of adjacent reflections introduces ambiguity into the assignment of integrated intensities, which are required to solve the phase problem. Nevertheless, significant advances have been made in this area^{7,8}. With the advent of a new generation of ultrahigh-resolution powder diffractometers, at the Rutherford-Appleton Laboratory, ILL Grenoble, and the Brookhaven National Light Source, a new era in structure determination is signalled, with the prospect that the structures of new materials that cannot be obtained as single crystals might be determined by powder methods. Here we illustrate the practicability of this approach with an *ab initio* study of the structure of ferric arsenate, FeAsO_4 .

The high-resolution powder diffractometer (HRPD) at the Rutherford-Appleton Laboratory's spallation neutron source, ISIS, has resolution ($\Delta d/d = 0.05\%$) that is approximately constant for all d -spacings and is superior to any previous neutron diffractometer. One of the many advantages of such high resolution is that it dramatically reduces the extent of peak overlap in the powder pattern, thus offering the prospect of obtaining a set of $I(hkl)$ values that is sufficiently extensive to permit the use of direct methods or Patterson techniques for structure solving.

The preparation of FeAsO_4 was first reported by Shafer *et al.*⁹. Subsequently, D'Yvoire¹⁰ indexed the X-ray powder pattern according to a monoclinic cell and assigned the space group $P2_1/n$. On the basis of infrared evidence and the facile transformation of the monoclinic modification to one with the CuSO_4 structure, it was suggested that the iron atom might be octahedrally coordinated¹⁰, but the structure has not been reported. The sample used for the neutron study was prepared from a nitric acid solution of As_2O_5 and $\text{Fe}(\text{NO}_3)_3 \cdot 9\text{H}_2\text{O}$. The solution was dried, the product ground, and fired for 1 week at 800°C . The homogeneity of the light brown powder was confirmed by analytical electron microscopy¹¹, and the morphology and particle size were examined by scanning electron microscopy. The crystallinities were rounded rather than well-faceted, and ranged in size from 0.1 to $1.0\ \mu\text{m}$. No impurity lines were observed in the X-ray powder pattern.

Time-of-flight neutron diffraction data were collected at room temperature on the diffractometer HRPD¹² at the spallation neutron source, ISIS. The sample, held in a slab-shaped vanadium container $25 \times 15 \times 10\ \text{mm}$, was placed at the high-

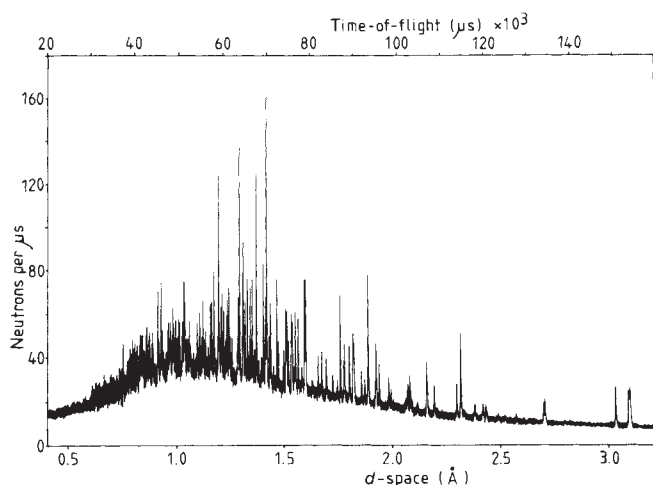


Fig. 1 The time-of-flight neutron diffraction pattern of FeAsO_4 before normalization. The shape of the background reflects the spectral distribution of the incident beam.

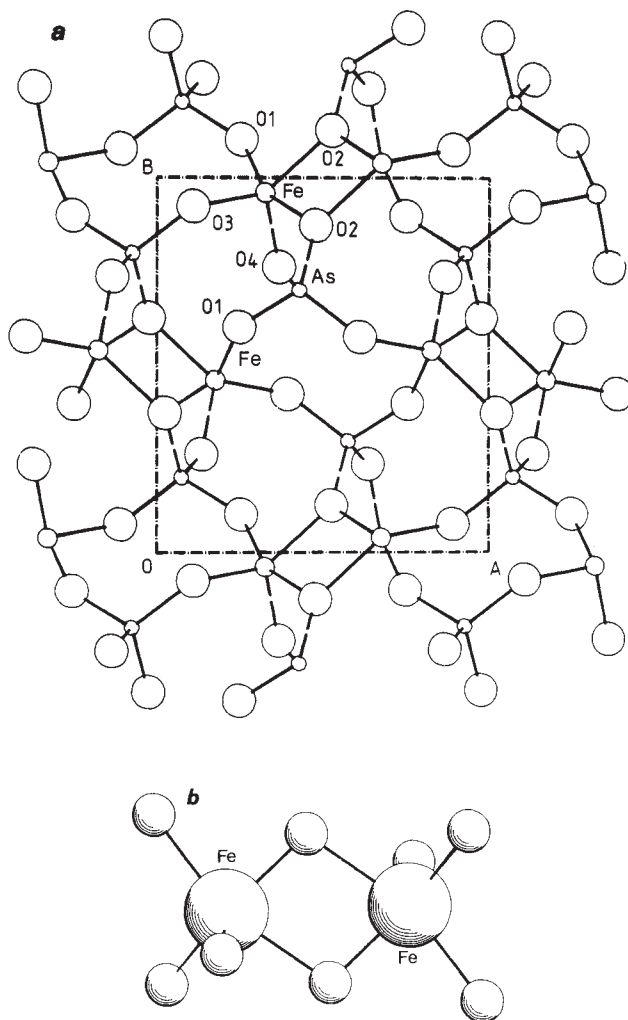


Fig. 2 *a*, A projection of the crystal structure of FeAsO_4 , down $[001]$. Broken bonds indicate connections to atoms that are obscured in the projection. *b*, The edge-sharing FeO_6 units in FeAsO_4 ; the principal axis of the trigonal bipyramid is $\text{O}(1)\text{--Fe--O}(2)$.

Table 1 α , Peak listing from direct methods solution

Peak no.	X	Y	Z	Height
1	0.178	0.449	0.758	2691
2	0.075	0.179	0.243	1960
3	0.210	0.080	0.413	1230
4	0.015	0.373	0.369	1199
5	0.898	0.046	0.178	1193
6	0.131	0.256	0.913	1104
7	0.562	0.587	0.726	591
8	0.385	0.617	0.887	590

<i>b</i> , Final atomic coordinates			
Atom	X	Y	Z
Fe	0.173 (2)*	0.462 (2)	0.763 (2)
As	0.073 (3)	0.202 (2)	0.223 (4)
O(1)	0.257 (3)	0.101 (3)	0.420 (4)
O(2)	0.027 (3)	0.377 (2)	0.384 (4)
O(3)	0.889 (4)	0.077 (3)	0.152 (4)
O(4)	0.125 (4)	0.267 (3)	0.939 (5)

139 reflections: triplets generated from top 66 reflections; negative quartets generated from top 67 reflections. Figure of merit, 4.0. Temperature factor, -3.8 (4).

* Estimated standard deviations are given in parentheses.

Table 2 Selected bond lengths and bond angles

Fe-O ₅ group		Bond lengths (Å)	Bond angle (deg.)
Fe-O(1)	1.914 (26)*	O(1)-Fe-O(2a)	159.3
-O(2a)	2.057 (25)	-O(2b)	87.2
-O(2b)	1.988 (25)	-O(3)	105.1
-O(3)	1.880 (28)	-O(4)	98.6
-O(4)	1.880 (30)	O(2a)-Fe-O(2b)	72.1
		-O(3)	88.7
		-O(4)	92.0
		O(2b)-Fe-O(3)	129.4
		-O(4)	120.4
		O(3)-Fe-O(4)	106.2

As-O ₄ group		Bond lengths (Å)	Bond angle (deg.)
As-O(1)	1.698 (30)	O(1)-As-O(2)	111.0
-O(2)	1.709 (28)	-O(3)	110.4
-O(3)	1.687 (31)	-O(4)	108.4
-O(4)	1.655 (33)	O(2)-As-O(3)	109.7
		-O(4)	105.5
		O(3)-As-O(4)	111.7

* Estimated standard deviations are given in parentheses.

resolution sample position (96-m primary flight path). The diffraction data were collected over a period of 3 days at a pulse repetition frequency of 3.125 Hz and a proton current of 0.8 μ A (at full design specification, a comparable run will take 45 min). The 40 spectra, obtained from the 20 concentric scintillator detectors ($170^\circ < 2\theta < 178^\circ$) of the two backscattering octants, were time-focused in software using the data reduction and manipulation program, GENIE¹³, to produce one summed spectrum (Fig. 1). This was then normalized by dividing the pattern by the incident spectral distribution at the sample, obtained by interpolation between readings from monitors placed before and after the sample.

The first step in the determination of the structure involved a check on the unit cell reported by D'Yvoire¹⁰, using an auto-indexing method¹⁴ with the d -spacings of 20 reflections. In the back-scattering geometry of HRPD, a few of the largest d -spacing reflections are inaccessible, and we were initially unable to obtain a satisfactory indexing of the pattern. However, with the addition of a single low-angle reflection from the X-ray powder diffraction pattern, a unique monoclinic cell ($a = 7.565$, $b = 8.077$, $c = 5.010$ Å, $\alpha = \gamma = 90^\circ$, $\beta = 104.5^\circ$) with a high figure of merit was speedily obtained. The cell constants were then refined by least-squares methods to give the final values: $a = 7.5636(1)$, $b = 8.0795(1)$, $c = 5.0117(1)$ Å, $\beta = 104.47(1)^\circ$, which are in excellent agreement with those reported by D'Yvoire¹⁰. A list of observed and calculated d -spacings is available from the authors. Examination of the systematic absences confirmed the space group $P2_1/n$.

Although variations in the peak shape made deconvolution by the Pawley method¹⁵ impossible, integrated Bragg intensities, $I(hkl)$, were obtained manually for 139 reflections in the range $1 < d < 3$ Å. This included approximately 60 weak reflections and a few overlapping peaks for which $I(hkl)$ values were apportioned arbitrarily. Structure factor amplitudes, $|F(hkl)|$, were then calculated in the normal way¹⁶ and were used as the input for a direct methods analysis using the programme MITHRIL¹⁷. Both triplets and negative quartets were calculated. The peak list from the calculation with the highest figure of merit (4.0) is shown in Table 1a. Analysis of the interpeak distances and angles confirmed that a chemically reasonable solution had been found, in which peaks 1 and 2 corresponded to Fe and As, respectively, and 3–6 to oxygen. Note that the peaks appear in order of their neutron scattering lengths and that there is a

substantial gap between 6 and 7; the latter peak indicates the level of noise in the E-map. A key factor in the successful use of direct methods to obtain trial coordinates, in this case, appears to be the use of quartets. The negative quartets harness the information content of the very weak reflections¹⁸, thus increasing the number of useful observations by a factor of 2. Calculations in which quartets were not used were unsuccessful.

The coordinates obtained from the direct methods analysis were then refined by using an integrated intensity least-squares procedure (GRILS), based on the Cambridge Crystallographic Subroutine Library¹⁹. A Rietveld profile analysis was precluded by problems with the description of the peak shape. No correction for extinction was necessary because of the small particle size ($< 1 \mu\text{m}$)²⁰. With 62 integrated intensities, the refinement converged to a final R -factor, $R(I) = 6.5\%$. The overall temperature factor is negative as a result of small systematic differences between detector and monitor efficiencies, but examination of the parameter correlation matrix showed that this had no effect on the precision and accuracy of the atomic coordinates. The refined values of the atomic coordinates and temperature factor are given in Table 1b, and the observed and calculated integrated intensities are available from the authors. The observed peak shapes are complex because, particularly at longer d -spacings, they are determined by sample effects, rather than instrumental resolution. There is a particularly rich variation of peak shape, not only with d -spacing but also with Miller index, resulting from both internal strain and anisotropic particle size broadening. Although these effects hinder profile analysis by the Rietveld method, they do indicate that, at high resolution, additional information may be obtained about macroscopic effects such as particle size, internal strain and stacking faults.

The crystal structure of FeAsO_4 is illustrated in Fig. 2a and the principal bond lengths and angles are given in Table 2. The compound is isostructural with LiAlH_4 (ref. 21), a result that was suggested at an early stage of the work by the similarity between their lattice parameters. We are unaware of any other compounds that adopt this structure. The most striking feature of the structure is the presence of five-coordinated iron, in marked contrast to the phosphorus analogue, FePO_4 , which adopts the berlinite structure with tetrahedrally coordinated cations²². The only other examples of five-coordinated iron in oxide systems are found in Fe_3PO_7 (ref. 23) and $\text{Fe}_4(\text{PO}_4)_2\text{O}$ (ref. 24), where, as in the present case, the stereochemistry is

approximately trigonal bipyramidal. The average Fe–O bond length, 1.944 Å, is in good agreement with the corresponding value in Fe_3PO_7 , 1.968 Å, and the As–O distances are in accord with previous studies on arsenates. The arrangement of FeO_5 and AsO_4 polyhedra is such that pairs of FeO_5 units share a common edge (Fig. 2b), as in $\text{Mg}_3(\text{PO}_4)_2$ (ref. 25). The Fe–Fe distance is, nevertheless, long (3.27 Å), suggesting that iron–iron repulsion is important; correspondingly, the O(2)–O(2) contact is rather short (2.38 Å).

We thank SERC for the provision of neutron facilities at ISIS, a studentship for R.J.B.J., and a PDRA for M.M.E. We especially thank all the staff at ISIS for their efforts in making this experiment possible, and Dr J. B. Forsyth and Professor A. J. Leadbetter for critical comments on the manuscript.

Received 14 October 1985; accepted 22 January 1986.

1. Hewat, A. W. *Nucl. Instrum. Meth.* **127**, 361–370 (1975).
2. Rietveld, H. M. *J. appl. Crystallogr.* **2**, 65–71 (1969).
3. Malmros, G. & Thomas, J. O. *J. appl. Crystallogr.* **10**, 7–11 (1977).

4. Cox, D. E., Hastings, J. B., Thomlinson, W. & Prewitt, C. *Nucl. Instrum. Meth.* **208**, 573–578 (1983).
5. Hathaway, B. J. & Hewat, A. W. *J. Solid State Chem.* **51**, 364–375 (1984).
6. Battle, P. D., Cheetham, A. K., Harrison, W. T. A., Pollard, N. J. & Faber, J. J. *J. Solid State Chem.* **58**, 221–225 (1985).
7. Berg, J.-E. & Werner, P.-E. *Z. Kristallogr.* **145**, 310–320 (1977).
8. Christensen, A. N., Lehmann, M. S. & Neilsen, M. *Aust. J. Phys.* **38**, 497–505 (1985).
9. Shafer, E. C., Shafer, M. W. & Roy, R. Z. *Kristallogr.* **108**, 263–275 (1956).
10. D'Yvoire, F. C. *r. hebdom. Séanc. Acad. Sci., Paris* **275C**, 949–951 (1972).
11. Cheetham, A. K. & Skarnulis, A. J. *Analyt. Chem.* **53**, 1060–1064 (1981).
12. Johnson, M. W. & David, W. I. F. *Rutherford-Appleton Lab. Rep. No. 85/112*.
13. GENIE *Rutherford-Appleton Lab. Rep. No. 86/015*.
14. Visser, J. W. *J. appl. Crystallogr.* **2**, 89–95 (1969).
15. Pawley, G. S. *J. appl. Crystallogr.* **14**, 357–361 (1981).
16. Buras, B. *Nucleonika* **8**, 259–266 (1963).
17. Gilmore, C. J. *J. appl. Crystallogr.* **17**, 42–46 (1984).
18. Schenk, H. *Computational Crystallography* (ed. Sayre, D.) 65–74 (Oxford University Press, 1982).
19. Matthewman, J. C., Thompson, P. & Brown, P. J. *J. appl. Crystallogr.* **15**, 167–173 (1982).
20. Sabine, T. M. *Aust. J. Phys.* **38**, 507–518 (1985).
21. Sklar, N. & Post, B. *Inorg. Chem.* **6**, 669–671 (1967).
22. Ng, N. M. & Calvo, C. *Can. J. Chem.* **53**, 2064–2067 (1975).
23. Modaresi, A., Courtois, A., Gerardin, R., Malaman, B. & Gleitzer, C. *J. Solid State Chem.* **47**, 245–255 (1983).
24. Bouchdoug, M., Courtois, A., Gerardin, R., Steinmetz, J. & Gleitzer, C. *J. Solid State Chem.* **42**, 149–157 (1982).
25. Nord, A. G. & Kierkegaard, P. *Acta chem. Scand.* **22**, 1466–1474 (1968).

The catastrophic failure of thermally tempered glass caused by small-particle impact

M. Munawar Chaudhri & Chen Liangyi*

Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, UK

One of the extensively used practical methods of strengthening soda–lime–silica glass against external stresses is to put its surface in a state of compression by thermal tempering or ion exchange^{1–3}. These processes place the subsurface regions under tension. The stress distribution in thermally tempered glass can be represented by a parabola, with the magnitude of the surface compression stress equal to twice the centre tension. A consequence of the bulk tensile stresses is that once a crack (generated by, for example, particle impact) reaches this tensile zone, the entire glass sample breaks up. In this situation, the energy required for producing cracks is elastically stored in the sample because of the strengthening procedure. There has been little work published on the manner in which a crack enters the tensile region, causing a catastrophic failure of the entire sample. Moreover, the velocities at which the cracks move through the different regions of a tempered sample are not known in detail, although a few measurements of the maximum crack velocities have been made^{3,4}. We now report the results of our high-speed photographic investigations, which provide clear answers to these questions. We also indicate the conditions under which the catastrophic failure of thermally tempered glass, caused by particle impact, is likely to occur.

The experiments were performed on soda–lime–silica blocks of $50 \times 40 \times 10$ mm and $50 \times 25 \times 12.5$ mm. The surfaces of the blocks were polished to an optical finish and a few of them were thermally tempered to produce a compressive stress of ~ 200 MPa on the surface, which decreased to zero at a depth of about one-fifth of the block thickness; beyond this plane the stress was tensile, reaching a maximum value of 100 MPa at the centre plane.

The test samples were subjected to normal-incidence impacts from tungsten carbide (WC) spheres of diameter 1 and 2 mm and sharp-cornered WC particles of irregular shape, of mass ~ 0.025 g, which were accelerated vertically to velocities of up to 250 m s^{-1} by an explosive gun system⁵. The impact-generated damage was photographed with a Beckman and Whitley model 189 rotating mirror camera at a framing speed of 10^6 s^{-1} ; at this

speed the exposure time per frame was $0.2 \mu\text{s}$. In a few experiments, a microscope was attached to the camera to provide a magnification of up to $\times 5$. The event was backlit with the flash of a xenon-filled FA5 tube and Ilford HP5 film (400 ASA) was usually used. In some experiments the stress patterns were recorded on Kodak VR 1000 colour film (1,000 ASA), using a photo-elastic arrangement for observing isochromatics⁶; with this arrangement the accuracy of the velocity measurement was 5%.

We found that spheres of 1-mm diameter and sharp-cornered particles with velocities of up to 250 m s^{-1} did not cause the catastrophic failure of the thermally tempered glass samples, though there was localized damage. However, WC spheres of 2-mm diameter with velocities of $\sim 150 \text{ m s}^{-1}$ did cause catastrophic failure. A high-speed photographic sequence is shown in Fig. 1, frames 1–6; the impact occurs on a 50×10 mm face and the camera views the event through a 50×40 mm face. In frame 1, the fringes in the glass represent the positions at which the difference between the two principal stresses, normal to the light beam, is constant. The fringe marked with an arrow indicates the position at which the magnitude of the stress is zero. The impact occurs in frame 2, as is evident from the generation of stress fringes below the contact. In the next frame, splinter cracks⁷ have appeared in the sample. The velocity of the cracks along the load axis is at least $1,000 \text{ m s}^{-1}$ between frames 2 and 3, and $1,500 \text{ m s}^{-1}$ between frames 4 and 5. A striking feature of this sequence is the 'tear-drop' shape of the stress fringes (frame 4), whereas, in annealed glass, the stress fringes are circular. Between frames 4 and 5 the cracks have entered the region of tensile stress and this leads to the initiation of the catastrophic failure of the sample. The cracks spread along the load axis and in the plane containing the load axis and the normal to the light beam (camera axis) at a uniform velocity of $1,500 \text{ m s}^{-1}$, thus giving a semicircular crack front. It is interesting, however, that the velocity of the cracks moving towards the surface of impact is much smaller (~ 200 – 300 m s^{-1}), possibly reducing to zero, as they approach it. Figure 1, frames 7–12 shows the manner in which the cracks spread through the sample. Note also that the cracks bifurcate, the distance between two bifurcating points along a crack path being ~ 0.5 – 1 mm; this dimension has been found to depend on the magnitude of the maximum tensile stress in the sample³.

The behaviour of the cracks moving in the direction of the impact and into the lower compression zone is better revealed in Fig. 2: here the impact occurs at the centre of a large face, $50 \times 25 \text{ mm}^2$, with the camera recording the damage through a $50 \times 12.5 \text{ mm}$ face. To reduce any effects of stress waves being reflected from the boundaries of the sample, another glass block, $100 \times 50 \times 25 \text{ mm}$, was placed on top of it. The impact occurs

* Present address: Xian Institute of Optics and Precision Mechanics, Xian, China.

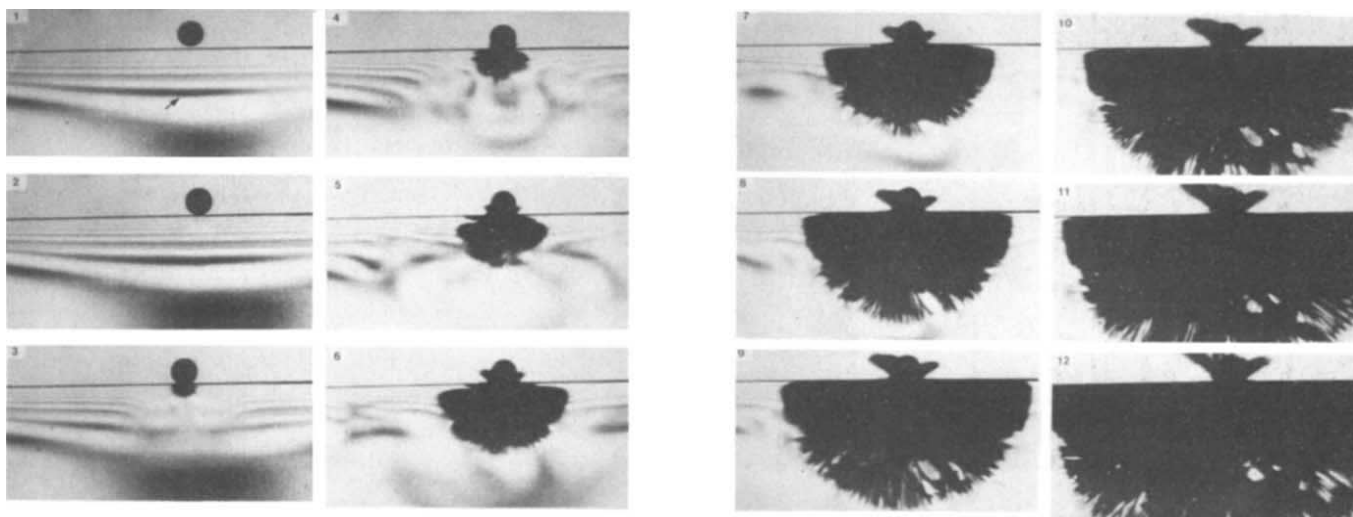


Fig. 1 High-speed photographic sequence showing the catastrophic failure of a thermally tempered glass block caused by the impact of a 2-mm-diameter tungsten carbide sphere at a velocity of 150 m s^{-1} . The fringes are isochromatics; the arrow points to the black fringe, which marks the region of zero stress. The catastrophic failure begins in frame 4. Note also the shape of the stress pattern in this frame which is quite different from that observed in an annealed block. Surface compressive stresses 200 MPa, inter-frame time $1 \mu\text{s}$, loading time of the impact cycle $3.5 \mu\text{s}$. 7-12, Continuation of the sequence shown in 1-6. Note the bifurcation of the cracks. (For clarity of presentation, all photographic sequences have been rotated by 180° about the normal to the page).

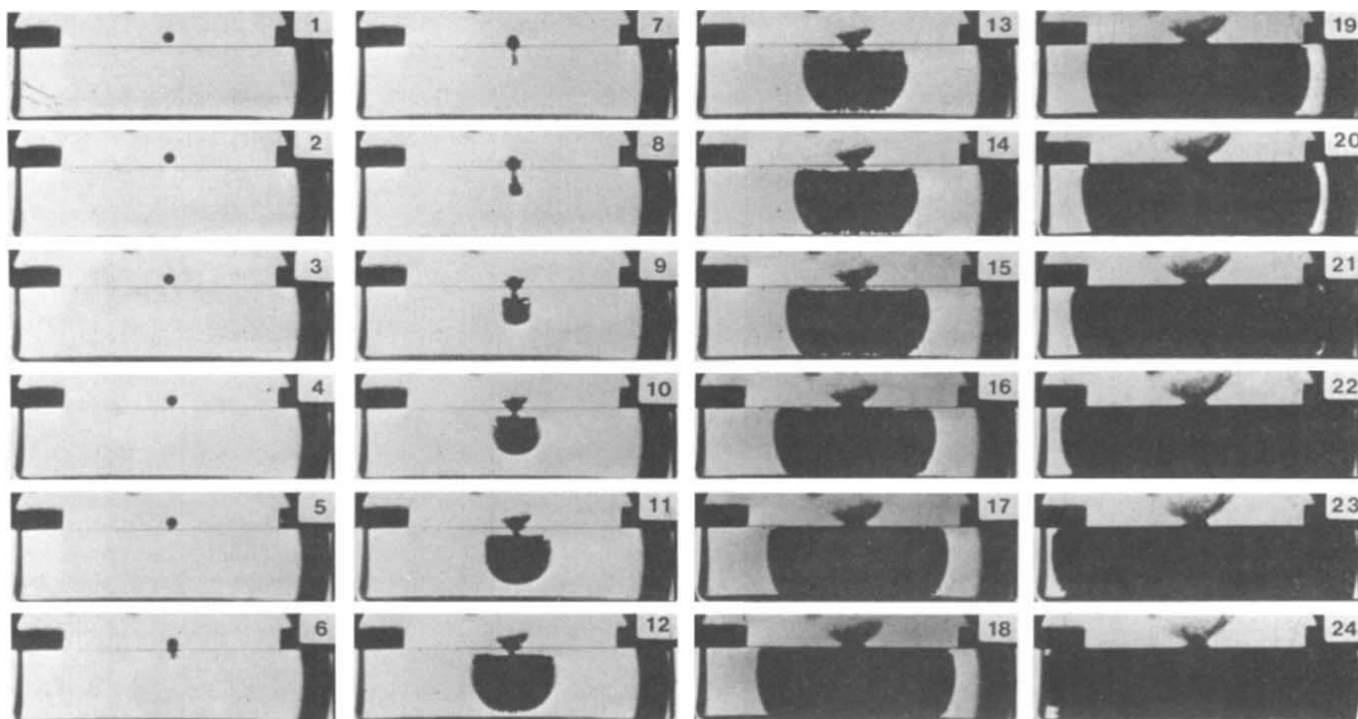


Fig. 2 High-speed photographic sequence showing the growth of the catastrophic failure of a thermally tempered glass block, caused by the impact of a 2-mm-diameter tungsten carbide sphere at 170 m s^{-1} . The camera views the damage through the thickness of the block and the sequence shows how the crack velocity falls from $1,800 \text{ m s}^{-1}$ in the tension zone to 200 m s^{-1} in the compression zone. Inter-frame time $1 \mu\text{s}$.

between frames 4 and 5 of Fig. 2, and the splinter cracks pass through the compression zone into the tensile zone at a velocity of $\sim 1,800 \text{ m s}^{-1}$ between frames 6 and 7; within the next microsecond the catastrophic failure begins about 2.5–3 mm below the surface of impact (frame 8). Initially, the cracks spreading to the left and to the right of the load axis accelerate, reaching a steady velocity of $\sim 1,800 \text{ m s}^{-1}$ within the next 2 microseconds. The velocity of the forward-moving cracks is $\sim 1,800 \text{ m s}^{-1}$ through the tensile zone of the specimen, but when the cracks reach the lower compression zone, the velocity falls dramatically to $\sim 200 \text{ m s}^{-1}$ (between frames 11 and 12) and it appears that,

by the time these forward-moving cracks reach the lower surface, their velocity has decreased further. The maximum velocity of $1,800 \text{ m s}^{-1}$ observed in the present work is slightly higher than the value of $\sim 1,500 \text{ m s}^{-1}$ reported earlier^{3,4}. The velocity of the cracks moving towards the surface of impact decreases from 900 m s^{-1} in frame 9 to $\sim 200 \text{ m s}^{-1}$ by frame 11, then it remains nearly constant for several microseconds as the cracks approach the surface.

Cracks (cone or splinter type⁷) which are initially inclined to the load axis, tend to bend over towards the impact surface during their propagation through the loading time of the impact;

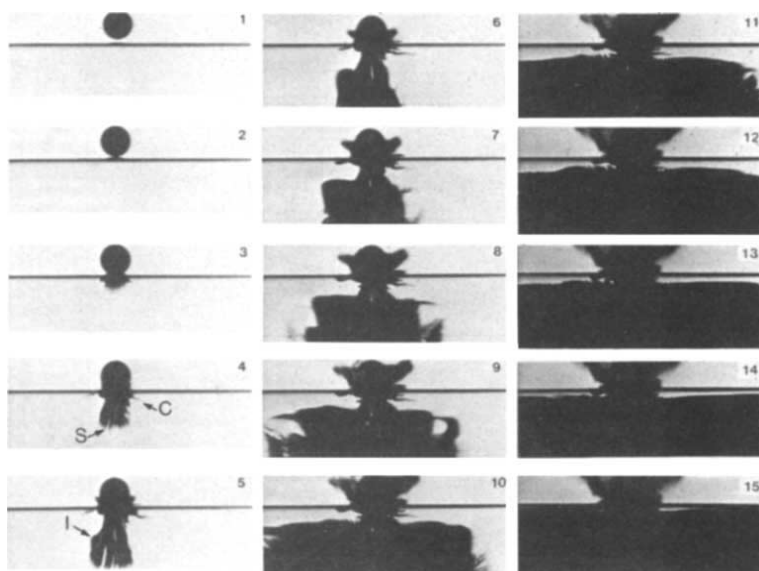


Fig. 3 High-speed photographic sequence, taken at high magnification, showing the details of the crack trajectories in a block of thermally toughened glass during the loading part of the impact of a 2-mm-diameter tungsten carbide sphere moving at 130 m s^{-1} . The crack shown at C (frame 4), which is initially formed at an angle to the load axis, bends over and runs approximately parallel to the impact surface because of the presence of the compressive stress. S (frame 4) indicates splinter cracks⁷ and I (frame 5) shows the position from which the catastrophic failure spreads. Inter-frame time $1 \mu\text{s}$. The camera records the damage through the thickness of the block.

this is clearly shown (Fig. 3) in a sequence taken at a higher magnification. Here the impact occurs at the centre of a $50 \times 25 \text{ mm}$ face between frames 2 and 3. Again, the effect of any stress waves was reduced as before. The cracking has begun in frame 3. Some of the cracks are splinter types and lie along the load axis (at S in frame 4) and some are inclined to the impact direction at a relatively large angle (C in frame 4). The velocity of the 'S' cracks between frames 3 and 4 is $\sim 1,800 \text{ m s}^{-1}$ and there is some indication of crack bifurcation. Between frames 5 and 6 crack 'C' bends over and then propagates approximately parallel to the impact surface. The bending over of the crack during the loading cycle is due to the presence of the compressive stresses; this is in accordance with the principal stress trajectories (M.M.C., unpublished work) for the case of a sphere loading on a surface having compressive stresses. The initiation of the catastrophic failure occurs at I in frame 5, and the rest of the frames show in detail the spreading of the cracks through the specimen. Again, the cracks approach the top surface at a relatively small velocity, $\sim 200 \text{ m s}^{-1}$, and possibly a thin surface layer is not crossed by them at all: evidence in support of this comes from recovered fragments of the sample.

Note that in none of our photographic sequences did we observe any effects on the crack pattern which could be attributed to the stress waves reflected from the sample boundaries, as observed by Field⁴ in a 6-mm-thick thermally tempered glass plate having an unspecified compressive stress.

Based on the photographic sequences presented here and others not shown, we suggest that, in order to cause the catastrophic failure of thermally tempered glass by small-particle impact, the loading conditions need to be such as to produce fast-moving splinter-type cracks, which will reach the tension zone within the loading cycle of the impact (most of the crack growth occurs during this time). The parameters of interest are: (1) the maximum impact load F_m , and (2) the loading time t of the collision. WC projectiles of 1- and 2-mm diameter travelling at velocities $>100 \text{ m s}^{-1}$ are likely to cause the plastic flow of the glass at the contact region. Then the para-

meters F_m and t are given by

$$F_m = \left(\frac{8P\rho}{3} \right)^{1/2} \pi R^2 v \quad (1)$$

$$t = \pi R \left(\frac{\rho}{6P} \right)^{1/2} \quad (2)$$

where ρ is the density of the projectile, which is 14.85 g cm^{-3} for WC; R its radius; v the impact velocity; and P the hardness of the tempered glass, which was measured to be 5.5 GPa . The calculated values of t for the 1- and 2-mm-diameter projectiles are 1 and $2 \mu\text{s}$ respectively; if crushing of the glass occurs, the effective value of P will be lower, causing an increase in t . For high values of F_m , the stress intensity factor driving the cracks will also be high, causing the cracks to propagate at their maximum velocity of $\sim 1,800 \text{ m s}^{-1}$ throughout the loading time, as observed in the case of the 2-mm projectiles. Then the predicted maximum length, L_m , of the impact-generated cracks is given by⁸

$$L_m \approx 1,800t \quad (3)$$

If L_m is greater than the thickness of the surface compressive layer, catastrophic failure can ensue. From equation (3), the predicted values of L_m for the 1- and 2-mm-diameter projectiles are 1.8 and 3.6 mm , respectively. As the thickness of the compressive layer in our samples is $2\text{--}3 \text{ mm}$, we expect the 2-mm projectiles to cause the catastrophic failure; this was indeed the case in our experiments.

Finally, the hertzian cone cracks are unlikely to be as effective in causing the catastrophic failure, because the 'skirts' of the cones will turn (depending on the magnitude of the surface stresses) and run approximately parallel to the surface of impact and thus may not propagate into the tension zone.

We thank Dr A. C. Woodward of Pilkington Brothers Ltd for providing some of the glass samples used in this work. C.L. thanks the Chinese Academy of Sciences and the Royal Society (London) for financial support.

Received 19 November 1985; accepted 22 January 1986.

1. Gardon, R. J. *non-cryst. Solids* **73**, 233–246 (1985).

2. Abou El Leil, M. M., Camaratta, F. A. & DiGenova, R. R. *J. Am. Ceram. Soc.* **68**, C18–C21 (1985).

3. Barsom, J. M. *J. Am. Ceram. Soc.* **51**, 75–78 (1968).

4. Field, J. E. *Times Sci. Rev. (Lond.)* **51**, 5–9 (Spring 1964).

5. Chaudhri, M. M., Knight, C. G. & Swain, M. V. *Proc. 12th Int. Cong. on High Speed Photography (Toronto 1976)*, SPIE **97**, 371–376 (Washington D.C., 1977).

6. Jessop, H. T. & Harris, F. C. *Photoelasticity* (Cleaver-Hume, London, 1949).

7. Knight, C. G., Swain, M. V. & Chaudhri, M. M. *J. Mater. Sci.* **12**, 1573–1586 (1977).

8. Chaudhri, M. M. & Walley, S. M. *Phil. Mag.* **A37**, 153–165 (1978).

Acceleration of solid-state β - α isomerization of the cobaloxime complex in mixed crystal containing the product

Akira Uchida, Yuji Ohashi* & Yoshio Sasada

Faculty of Science, Tokyo Institute of Technology, Nagatsuta, Midori-ku, Yokohama 227, Japan

It has been reported¹ that when β -cyanoethyl cobaloximes with various base ligands (cobaloxime = 2,3-butanedionedioximatecobalt(III)) are exposed to visible light in the solid state, β - α isomerization of the cyanoethyl group occurs¹. During a detailed study of one such reaction, the β - α isomerization of (β -cyanoethyl)(3-methylpyridine)cobaloxime, we found that some batches of synthesized material were isomerized abnormally quickly. Using crystallography, we show here that the anomalous crystals are mixtures of the original material and the reaction product with an α -cyanoethyl group. To our knowledge, this is the first example of an 'autocatalytic' solid-state reaction and of the possible artificial control of a reaction achieved by modifying crystals, although the use of mixed crystal has been reported for engineering the solid-state reaction of benzylbenzylidenecyclopropanones². The explanation seems to be that there is more room in the mixed crystals for the movement of the atoms of the cyanoethyl group involved in the β - α isomerization, which is consistent with the conclusions of our earlier studies of crystalline-state racemization³⁻⁵.

Very little is known about the reaction mechanism of inorganic complexes and organometallic compounds in the solid state⁶. Serial studies on crystalline-state racemization³⁻⁵ have revealed a correlation between the amount of room that the atoms of the reactive group have to move in—the reaction cavity—and the reaction rate for some cobaloxime complexes. We have extended our research to the β - α isomerization of (β -cyanoethyl)(3-methylpyridine)cobaloxime. The complex was synthesized by absorption of hydrogen in methanol solution of CoCl_2 , 2,3-butanedionedioxime, acrylonitrile and 3-methylpyridine. Slow cooling of the solution from about 60 °C produced red-brown crystals. The crystals are monoclinic, space group $P2_1/a$ with unit cell dimensions of $a = 23.742(5)$, $b = 9.496(2)$, $c = 8.819(1)$ Å and $\beta = 94.59(2)^\circ$. Structure analysis was undertaken with MULTAN78 (ref. 7) to give the final R value of 0.059 for 3,958 reflections with $|F_o| > 3\sigma$.

The crystals obtained from another synthesis batch are isomorphous to the above crystals; unit cell dimensions are $a = 23.764(5)$, $b = 9.592(2)$, $c = 8.843(2)$ Å and $\beta = 93.90(2)^\circ$. The change in cell dimensions is anisotropic; the maximum variation is in b , and the increase in volume of the cell is ~ 29 Å³. On the difference map, two extra peaks appeared, which were assigned to the methyl groups of α -cyanoethyl groups with R and S configurations. Constrained refinement for β - and α -cyanoethyl groups gave the occupancy factors of 0.66 and 0.34, respectively. The final R value was 0.066 for 3,110 reflections. Mixed crystal formation was further confirmed by the infrared spectra of the crystal. Figure 1 shows the molecule in the mixed crystal, the hatched ellipsoids corresponding to the atoms in the α -cyanoethyl group. Although the shapes of β - and α -cyanoethyl groups seemed quite different at first sight, the structure determined implies that mixed crystal formation is a reasonable hypothesis. Bond lengths and angles are in good agreement between the pure and mixed crystals. Conformations of the β -cyanoethyl group and 3-methylpyridine ligand do not change significantly. Figure 2 compares the structures of the two crystals.

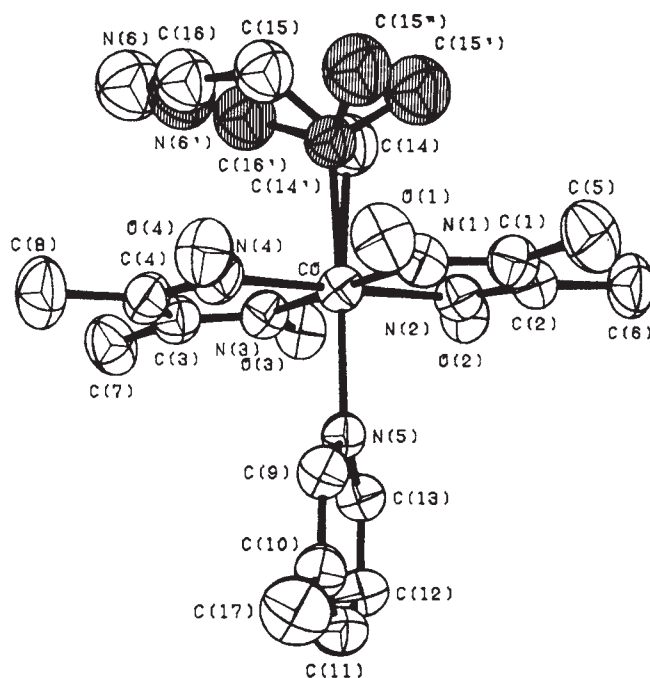


Fig. 1 Structure of the molecule in the mixed crystal. Hatched ellipsoids are the atoms of the disordered racemic α -cyanoethyl group, whose overall occupancy is 0.34. The methyl carbon atoms in the α -cyanoethyl groups with S and R configurations are $C(15')$ and $C(15'')$, respectively.

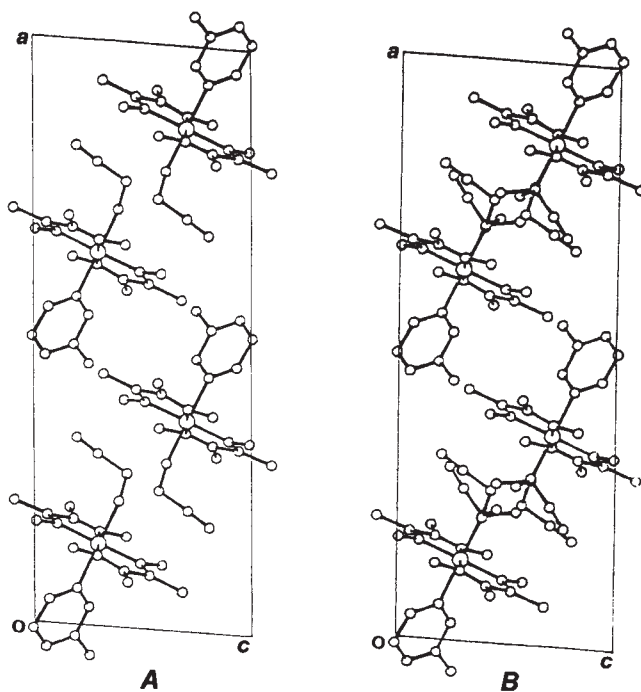


Fig. 2 Structures of the pure (A) and mixed (B) crystals.

An experiment to determine isomerization by visible light was carried out using powdered crystals. The degree of conversion was estimated by the decrease in absorbance of the $2,250\text{-cm}^{-1}$ band which corresponds to the $\text{C}\equiv\text{N}$ stretching of the β -cyanoethyl group. Infrared spectra were taken with Hitachi spectrophotometer 260-10. Figure 3 plots the conversion against the exposure time for the two crystals. The first-order best-fit

* Present address: Faculty of Science, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112, Japan.

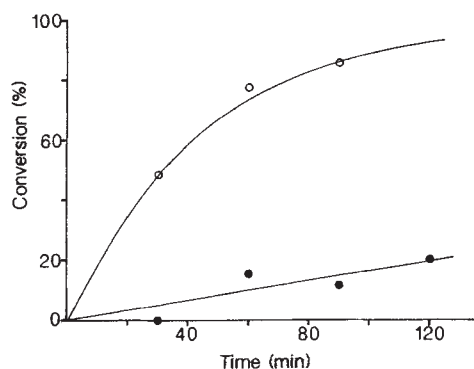


Fig. 3 β - α Isomerization of the cyanoethyl group. KBr disk containing 1–3 mg of powdered crystals were irradiated by 500-W mercury lamp at 1 m distance. To avoid oxidation by air, the disk was put in a transparent vessel filled with argon. Conversion was estimated from the decrease in absorbance of the $2,250\text{-cm}^{-1}$ band.

kinetics curves were drawn through the points. The rate constants were 0.3×10^{-4} and $3.6 \times 10^{-4} \text{ s}^{-1}$ for the pure and mixed crystals, respectively. Mixed crystal formation greatly accelerates the β - α isomerization, the rate constant being 12 times larger than for pure crystals. The acceleration can be understood by drawing the reaction cavity for the β -cyanoethyl group, as defined previously³. Figure 4 shows that the cavity in the mixed crystal becomes slightly thicker in the neighbourhood of the methyl group of the α -cyanoethyl group. Its volume is $11.89 \text{ \AA}^3$, larger by $1.5 \text{ \AA}^3$ than that of the pure crystal ($10.36 \text{ \AA}^3$).

The present finding demonstrates the possibility of controlling the reaction rate by artificial modification of the crystalline field. A more systematic search for mixed crystal formation would be useful as the basis for designing the reactive crystal².

This work was supported by a Grant-in-Aid for Science Research from the Ministry of Education, Science and Culture, Japan (59470126).

Received 14 October; accepted 30 December 1985.

1. Ohgo, Y. & Takeuchi, S. *JCS chem. Commun.* 21–23 (1985).
2. Theocharis, C. R., Desiraju, G. R. & Jones, W. J. *Am. chem. Soc.* **106**, 3606–3609 (1984).
3. Ohashi, Y., Uchida, A., Sasada, Y. & Ohgo, Y. *Acta crystallogr. B* **39**, 54–61 (1983).
4. Ohashi, Y., Yanagi, K., Kurihara, T., Sasada, Y. & Ohgo, Y. *J. Am. chem. Soc.* **103**, 5805–5812 (1981).
5. Tomotake, Y. *et al. Isr. J. Chem.* **25**, 327–333 (1985).
6. O'Brien, P. *Polyhedron* **2**, 233–243 (1983).
7. Main, P. *et al. MULTAN78, A system of Computer Programs for the Automatic Solution of Crystal Structures from X-Ray Diffraction Data* (Universities of York, UK, and Louvain, Belgium, 1978).

Controlled-source electromagnetic sounding of the oceanic lithosphere

C. S. Cox*, S. C. Constable*, A. D. Chave†
& S. C. Webb‡

* Scripps Institution of Oceanography and † Institute of Geophysics and Planetary Physics, University of California, San Diego, La Jolla, California 92093, USA

‡ Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

The attenuation of ionospheric signals in the frequency range 0.06–24 Hz by sea water effectively precludes using the magnetotelluric method to study the electrical structure of the upper oceanic lithosphere. We have carried out a dipole–dipole electromagnetic sounding in the North Pacific by injecting electromagnetic signals into the ocean and sea bed. The crust at the site is 25 Myr old and has a thin sediment cover. The technique, similar to that used in earlier work^{1,2}, involves dragging a horizontal dipole antenna along the sea floor. The electric fields that propagated through the resistive basement were detected by seafloor receivers at ranges of 10–65 km. As the ambient electric field is very small (varying from $10^{-18} \text{ V}^2 \text{ m}^{-2} \text{ Hz}^{-1}$ at 0.1 Hz to $10^{-24} \text{ V}^2 \text{ m}^{-2} \text{ Hz}^{-1}$ above 1 Hz; ref. 3), the controlled-source signals could be easily monitored. Our data are consistent with a simple one-dimensional Earth model consisting of a 3–7-km-thick crustal layer of moderate conductivity ($\sim 0.001 \text{ S m}^{-1}$) underlain by a thicker region of very low conductivity ($< 2 \times 10^{-5} \text{ S m}^{-1}$). The results suggest an upper mantle water content of at most 0.1% by volume.

The experiment was performed by towing a transmitting antenna consisting of a 600-m section of insulated armoured cable coupled to the ocean by 15 m of bared cable at the ends. The antenna was energized using a waveform synthesizer supplied with 60-Hz power from a research ship's generators through a towing cable. Rectified half-cycles of the 60-Hz power were modulated by bipolar switching to produce a variable-frequency waveform having equal first and third harmonics. The

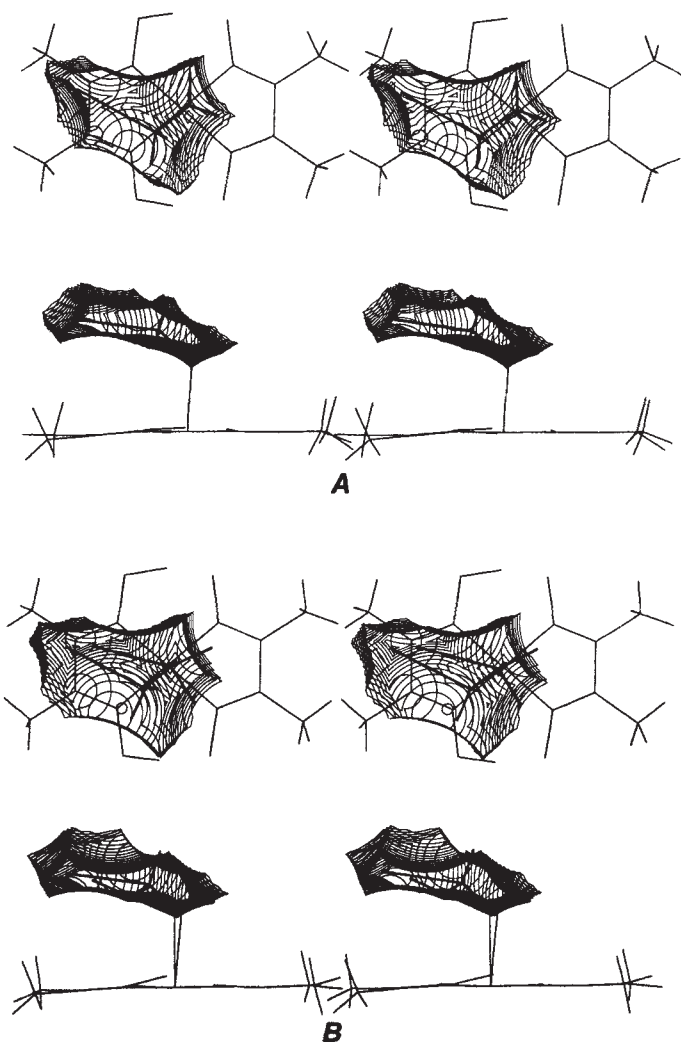


Fig. 4 Cavities for cyanoethyl groups in the pure (A) and mixed (B) crystals (stereo pairs). They were defined as the space limited by a concave surface of the spheres of the surrounding atoms, whose radii were taken to be $1.2 \text{ \AA}$ greater than the van der Waals radii of the corresponding atoms³. In the cavity, the atomic centres of the reactive group can be located freely. Upper, view along the normal to the cobaloxime plane; lower, side view. Contours were drawn in sections separated by $0.1 \text{ \AA}$.

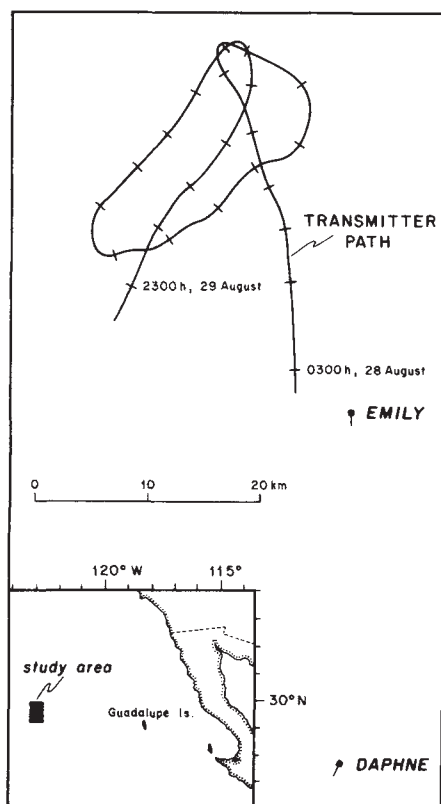


Fig. 1 The inset shows the location of the larger map of the study area, at 28° N, 122° W. The path of the ship during towing of the transmitter was derived by smoothing navigational data collected every 15 min. The cross-marks show ship positions at 2-h intervals. Emily and Daphne are the two electrical recorders. The age of the crust is 25 Myr (magnetic anomaly 6E, after ref. 11).

current amplitude at both frequencies was 65 A. The receivers⁴ were self-contained recording voltmeters which measured a single component of the horizontal electric field averaged over a 600-m antenna of insulated wire. The antennas were coupled to the ocean by low-noise, silver/silver chloride electrodes. The receivers were programmed to stack the received signal synchronously with the transmitted waveform, and the source frequency was stepped by a factor of 2 at 30-min intervals.

The transmitting antenna was towed over the sea floor so that it remained in continuous contact with the sea bed. Owing to imperfections in navigation and the effect of varying winds and currents on the towing ship and cable, it is not certain that the antenna stretched out in a straight line, but its contact with the sea floor was monitored continuously with acoustic methods. The path of the ship during the tow is illustrated in Fig. 1, which also gives the locations and orientations of the two receivers which have yielded the best data.

The propagation of electromagnetic signals at the transmitted frequencies in conducting sea water and rocks is essentially a diffusive process because the conduction current far exceeds the displacement current. The attenuation of an electromagnetic disturbance with range is very rapid in sea water, and the detection of the source at the long distances used in this experiment implies propagation through the rocks below the ocean floor. The variation of the amplitude of the electric field with range and frequency provides information on the sub-seafloor electrical conductivity as a function of depth. The theory is well-developed for one-dimensional media, where the electromagnetic fields separate into two independent modes associated with the vertical electric and magnetic fields respectively⁵.

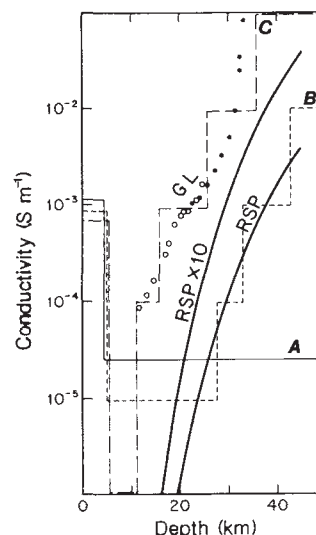


Fig. 2 Three acceptable models of conductivity, compared with laboratory data. The laboratory measurements have been interpreted as conductivity versus depth below the sea floor by assuming a uniform thermal gradient of 33 K km⁻¹ (ref. 12). GL, the conductivity of a garnet lherzolite at atmospheric pressure¹²; O, measurements of a natural rock sample; ●, measurements from a powdered, melted and recrystallized sample. Both measurements were made while the temperature was rising. RSP is the conductivity of a single crystal of Red Sea peridot¹⁴ including a pressure correction, and RSP×10 is the 10 times larger conductivity suggested by Shankland and Waff¹⁵ for mantle conductivity. Model A is the best-fitting two-layer model. Although the second layer extends to infinite depth, the data are insensitive to structure deeper than 35–40 km. In models B and C the deeper conductive layers have been constrained to fit RSP and GL laboratory data respectively, but the most resistive and the shallowest layers have been optimized during the model-fitting.

At the highest frequencies transmitted, the measured electric fields were surprisingly strong and had signal-to-noise ratios >10 (after stacking) for all separations of the transmitter and receiver. The signals at the lowest frequencies were weaker and more difficult to detect because of a higher noise level. Frequency-dependent polarization of either the transmitter or receiver electrodes would explain this type of variation. This hypothesis has been rejected on the basis of laboratory studies of the electrodes. Furthermore, the received signals were so much weaker than the broadband background noise level that polarization losses would be associated mainly with the low-frequency components of the noise. An observed independence of apparent conductivity with range despite a large decrease in amplitude of the electric field implies that any nonlinear amplitude-dependent bias is small. However, failure to stretch either the transmitter or receiver antennas to their full length would produce a systematic downward bias in the observed fields. Several independent cross-receiver checks and the reproducibility of the signal amplitude, especially at short and moderate ranges, suggests that these biases are unimportant. An earlier experiment in 1983 also gave results consistent with this experiment.

Layered structures which fit the data adequately were constructed using trial and error and Marquardt fitting techniques⁶. Model-fitting procedures of this type do not have unique solutions, and a variety of models have been constructed which fit the data equally well. Three such models are shown in Fig. 2. Although it is difficult to give bounds for all acceptable models, common characteristics of the models studied are evident. The first characteristic is a well-resolved upper layer, ~5 km thick,

of moderate conductivity (10^{-3} S m^{-1}), presumably representing cool, fractured and water-saturated crust. (The seismically estimated crustal thickness at $27^{\circ}24' \text{ N}$, $212^{\circ}35' \text{ W}$ is 7.4 km^7 .) The inclusion of a more conductive layer of sediment and fractured basalt at the surface has no apparent effect on the model fits.

A second characteristic is that as the deep structure below 10–20 km is made more conductive, the conductivity of the sub-crustal zone below 5 km must decrease. In the extreme case, where no increase in conductivity at depth is included (model A), the sub-crustal zone attains a maximum conductivity of $2 \times 10^{-5} \text{ S m}^{-1}$. Model A, however, is impossibly resistive at depths $>30 \text{ km}$. To illustrate this, laboratory conductivity data for upper mantle materials have been plotted in Fig. 2. The two sets of laboratory data represent fertile (GL) and completely depleted (RSP) models of mantle conductivity in the sense that partial melting of garnet lherzolite yields basalt while the pure olivine of peridot cannot. Both sets show the conductivity increase associated with the increase of temperature with depth in the mantle. As models of the conductivity in the subsolidus upper mantle, both sets of data suffer from difficulties. The defect structures in the GL sample were probably not in equilibrium at each temperature and there may have been some thermally driven disruption of the grain-to-grain cohesion. The single-crystal measurement of RSP is probably an underestimate of the mantle conductivity because grain boundary effects and impurities tend to increase the conductivity of silicates. Both of these models represent completely dry rocks. If water is present, the conductivity in the comparatively cool parts of the uppermost mantle will be increased greatly by the mobility of ions in the intergranular fluid and in adsorbed films on the grain boundaries⁸.

The laboratory data may be used to constrain the deep structure of the electrical models fitting the electromagnetic data (models B and C). If the RSP curve is taken to be an underestimate of mantle conductivity, we see that the conductivity of the sub-crustal zone is now constrained to be $<10^{-5} \text{ S m}^{-1}$. We may use this constraint to place a rough upper limit on the volume fraction of water present in the upper mantle.

For this discussion we assume that the bulk conductivity of the rocks containing intergranular fluids follows Archie's law in the form

$$\sigma = \sigma_f p^2$$

where σ_f is the conductivity of the fluid and p is the volume fraction of interconnected fluid spaces. We neglect entirely surface effects of the intergranular fluid. The conductivity of 0.1 M NaCl solutions is $\sigma_f = 5.2 \pm 0.3 \text{ S m}^{-1}$ in the relevant temperature and pressure ranges ($300\text{--}600^\circ \text{C}$ and $0.3\text{--}0.6 \text{ GPa}$)⁹. The upper limit on the volume fraction of water is then $p = 1.4 \times 10^{-3}$. For more saline water, the upper limit on water content is even lower. If the vapour pressure is lithostatic, as we have assumed, the upper limit on the mass fraction of interconnected water is then $\sim 4 \times 10^{-4}$. These numbers show that no appreciable sea water has leaked into the upper mantle, in agreement with indications that serpentinization of mantle rocks in ophiolites occurs after obduction¹⁰, and that juvenile

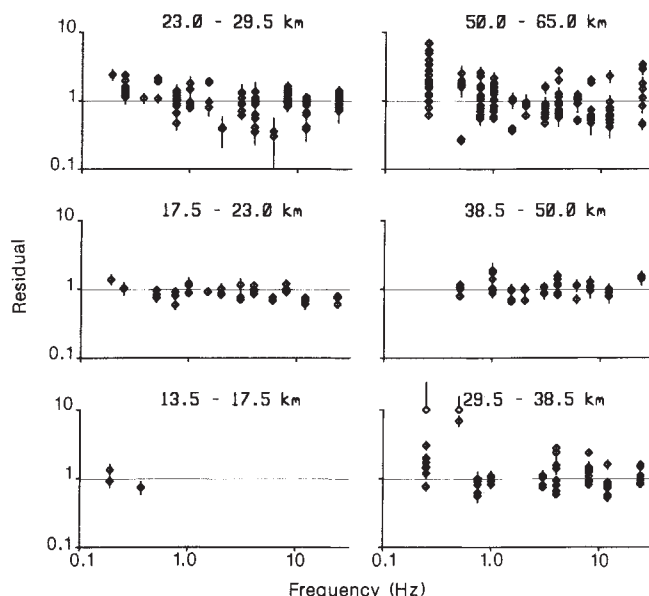


Fig. 3 The residuals (amplitude of the measured electric field component divided by the amplitude of the model prediction) are shown as a function of frequency for six different range bins for the two-layer model shown as line A in Fig. 2. The vertical bars represent standard errors in measuring the electric fields. The two data plotted at 10 Hz for a range of 29.5–38.5 km lie off the plot. The data vary over 4 orders of magnitude, from 10^{-17} to $10^{-13} \text{ V A}^{-1} \text{ m}^{-2}$, although the residuals show that the model fits most of the data to within a factor of 2.

water, if present in the mantle as a whole, must have been swept out during lithospheric generation.

Considerable scatter of the observations about the best-fitting model predictions is observed (Fig. 3). This is not entirely the result of errors caused by a rising noise level, as the signal-to-noise ratio remains large, even allowing for errors in the orientation of the transmitter and receiver antennas. Furthermore, the scatter is not evenly distributed across the range bins. There is evidence that the residuals are a function of transmitter position, suggesting that irregular conductivity structure in the region of the transmitter tow is causing the scatter. However, the fact that there is no systematic structure left in the residuals suggests that the one-dimensional modelling gives a reasonable picture of the gross lithospheric structure.

This work was supported by the NSF and the Defense Advanced Research Projects Agency through the Naval Ocean Research and Development Activity. Technical assistance was provided by T. Deaton, J. Lyons and J. Lemire. Captains T. Desjardins of the RV *New Horizon* and T. Beatty of the RV *Ellen B. Scripps*, their crews and G. Hargreaves were exceptionally helpful during the experiments.

Received 26 July; accepted 30 December 1985.

1. Spiess, F. N. *et al. Science* **207**, 1421–1433 (1980).
2. Young, P. D. & Cox, C. S. *Geophys. Res. Lett.* **8**, 1043–1046 (1981).
3. Webb, S. C. & Cox, C. S. *Geophys. Res. Lett.* **11**, 967–970 (1984).
4. Webb, S. C., Constable, S. C., Cox, C. S. & Deaton, T. K. *J. Geomagn. Geoelectr.* (in the press).
5. Chave, A. D. & Cox, C. S. *J. geophys. Res.* **87**, 5327–5338 (1982).
6. Marquardt, D. W. *J. Soc. ind. appl. Math.* **11**, 431–441 (1963).
7. Shor, G. G. Jr, Menard, H. W. & Raitt, R. W. in *The Sea* Vol. 4, Pt 2 (ed. Maxwell, A. E.) 3–27 (Wiley, New York, 1970).

8. Brace, W. F. *Am. geophys. Un. Monogr.* **14**, 243–260 (1971).
9. Quist, A. S. & Marshall, W. L. *J. phys. Chem.* **71**, 684–703 (1968).
10. Gregory, R. T. & Taylor, H. P. *J. geophys. Res.* **86**, 2737–2755 (1981).
11. Klitgord, K. G. & Mammereckx, J. *J. geophys. Res.* **87**, 6725–6750 (1982).
12. Parsons, B. & Slater, J. G. *J. geophys. Res.* **82**, 803–821 (1977).
13. Rai, C. S. & Manghnani, M. *Phys. Earth planet. Inter.* **17**, 6–13 (1978).
14. Duba, A., Heard, H. C. & Schock, R. N. *J. geophys. Res.* **79**, 1667–1673 (1974).
15. Shankland, B. & Waff, H. S. *J. geophys. Res.* **82**, 5409–5417 (1977).

Terrestrial analogues of the surface rocks of Mars?

Roger G. Burns

Department of Earth, Atmospheric and Planetary Sciences,
Massachusetts Institute of Technology, Cambridge,
Massachusetts 02139, USA

The pathways of chemical weathering processes on Earth, and the virtual absence of them on the Moon, are well characterized because we have direct access to samples of primary igneous rocks, their upper mantle sources and derivative sedimentary rocks and soil forming the regolith. Tracing chemical changes on the surface of Mars, however, is more difficult because of limited access to samples from Mars. Terrestrial rocks and minerals are continually being sought which might serve as analogues for martian materials. Evidence presented here suggests that hisingerite, iddingsite and several ferric sulphate minerals derived from deuteritic alteration of iron-rich basalts may provide clues to the evolutionary history of the surface of Mars.

The red colour of the surface of Mars has long been attributed to Fe³⁺-bearing minerals in the regolith, implying that oxidizing conditions have prevailed. However, the apparent enrichment of iron in martian regolith, its mineralogy and the nature of the unweathered parent rocks remain poorly defined. The compositions of parent magma and upper mantle source rocks have been deduced^{1,2} from various pieces of evidence, including Earth-based and remote-sensed observations and *in situ* Viking Lander measurements of the martian regolith³. The mantle of Mars (Table 1, columns 1 and 2) is believed to be considerably more iron-rich than the Earth's mantle^{1,2}, with estimated Mg/(Mg + Fe) ratios ranging from 0.67 to 0.77, compared with 0.85 for the Earth. It has also been suggested² that the high FeO content of the martian mantle and the oxidizing effect of late-accreting, volatile rich material allows significant amounts of NiO and Fe₂O₃ to exist in the interior of Mars. Partial melting of the proposed martian mantle would lead to an iron-rich basaltic magma¹ (Table 1, column 3) which could be olivine-rich and/or pyroxene-rich. The normative mineralogy of such mafic or ultramafic rocks indicates the presence of fayalitic olivines and Fe-rich pyroxenes^{1,4} which are highly vulnerable to chemical weathering⁵ on the surface of Mars. Pyrrhotite was suggested to be a sulphide accessory mineral¹.

The Viking X-ray fluorescence (XRF) measurements^{6,7} (Table 1, column 5) indicated that the regolith of Mars is significantly

enriched in Fe, S and Cl, and depleted in Al compared with chemically weathered basalts on Earth. The bulk composition of the surface fines was suggested to match a mixture of Fe³⁺-rich smectite clay, Mg sulphate and iron(III) oxide phases apparently present in a caliche-like duricrust, the nature of which indicates the intervention of brine during surface weathering of Mars.⁸ The identity and origin of the clay silicate phases has been discussed extensively⁹, and the iron montmorillonite (smectite, nontronite) identified in subaqueous palagonite, in soils above weathered basalt, and in altered impact or volcanic glasses on Earth is generally accepted to be the most probable phyllosilicate mineral in the martian regolith¹⁰⁻¹².

Compositional similarities between the rare SNG (shergottites, nakhlites, chassignites) meteorites and martian fines have been noted^{5,13}. Indeed, their textures and mineralogy, consisting of volatile-rich mafic cumulates containing fayalitic olivines, iron-rich pyroxenes, and accessory pyrrhotite or troilite¹⁴, contribute to the common belief that SNG meteorites are derived from Mars¹⁵. The question arises as to whether other primary igneous rock types and secondary minerals found on Earth might serve as analogues of martian materials and aid in a better understanding of geochemical processes on Mars.

The iron-rich basaltic magmas suggested to have flowed onto the surface of Mars may have terrestrial analogues in the iron-rich gabbros and diabases occurring in the Skaergaard (Greenland) intrusion, Beaver Bay (Minnesota) complex and New Almalfi (South Africa) intrusive sheet¹⁶. Of particular interest is the Beaver Bay diabase which is intrusive into the Precambrian Keeweenaw volcanic series above the Duluth gabbro^{16,17}. Fayalitic olivines and iron-rich pyroxenes contribute to the mineralogy of the diabase, the bulk composition of which is summarized in Table 1 (column 4). This diabase has notable similarities to the 20% partial melt composition of the martian mantle proposed by McGetchin and Smyth¹ (Table 1, column 3), particularly with regard to the high Fe and Al contents.

A common alteration product of Beaver Bay iron-rich gabbroic rocks is hisingerite, a very poorly crystalline hydrous iron silicate formed by deuteritic and late-state hydrothermal alteration of pyroxenes and olivine¹⁸. In massive form, hisingerite has a pitchy luster, conchoidal fracture and pulverizes to an orange-brown powder¹⁹ resembling the colour of martian regolith. Some broad, diffuse lines in X-ray diffraction patterns of hisingerite resemble those of nontronite²⁰, the phyllosilicate commonly believed to occur on Mars. Moreover, a chemical analysis of the Beaver Bay hisingerite¹⁸ shown in Table 1 (column 6) bears some significant resemblances to the Viking XRF analyses of the martian regolith⁷ (Table 1, column 5), including high Fe, constant Si and depleted Al contents relative

Table 1 Chemical compositions of terrestrial materials relevant to Mars (wt %)

	1	2	3	4	5	6	7
SiO ₂	40.04	41.60	43.48	45.87	44, 43	42.23	41.20
TiO ₂	0.63	0.33	—	4.65	0.62, 0.54	0.13	0.14
Al ₂ O ₃	3.14	6.39	11.25	10.66	7.3, (7)	3.65	3.81
Fe ₂ O ₃	0.41	—	3.00	4.11	17.5, 17.3	23.28	35.42
FeO	18.48	15.85	11.38	15.59	—	5.90	0.38
MnO	0.12	0.15	—	0.29	—	0.92	0.05
MgO	33.22	29.78	17.32	3.50	6, (6)	5.21	6.85
CaO	2.73	5.16	9.54	8.28	5.7, 5.7	2.52	2.35
Na ₂ O	0.51	0.10	1.77	2.84	—	0.27	0.12
K ₂ O	0.12	0.01	0.40	0.96	<0.5, <0.5	0.07	0.10
Others	0.61	0.65	—	1.70	7.5, 8.3	*	†
Total:	100.01	100.02	98.14	100.34	91, 90	99.63	99.54

(1) Mars mantle, based on McGetchin and Smyth¹, from Morgan and Anders (ref. 2). Includes: Cr₂O₃, 0.63; NiO, 0.18; P₂O₅, 0.05. (2) Mars mantle, from ref. 2. (3) 20% Partial melt of Mars mantle, from ref. 1. (4) Iron-rich diabase, Beaver Bay, Minnesota¹⁶. Analysis includes: P₂O₅, 1.70. (5) Mars regolith, average analyses at Viking Lander 1 (Chryse) and 2 (Utopia) sites, from Clark *et al.*⁷. Includes: Cl, 0.8, 0.4; SO₃, 6.7, 7.9; P + Mn + Na, 2%. (6) Hisingerite, associated with olivine gabbro, Beaver Bay, Minnesota¹⁸. (7) Iddingsite, average of 10 analyses, from Wilshire²².

* Includes H₂O(+), 5.82; H₂O(−), 13.54.

† Includes H₂O(+), 9.30; H₂O(−).

Table 2 Selected examples of ferric sulphates and related minerals

Mineral	Formula	Colour
Coquimbite	$\text{Fe}_2(\text{SO}_4)_3 \cdot 9\text{H}_2\text{O}$	Violet-amethyst-yellow
Hohmannite	$\text{Fe}_2(\text{SO}_4)_2(\text{OH})_2 \cdot 7\text{H}_2\text{O}$	Chestnut brown-burnt orange
Fibroferrite	$\text{FeSO}_4(\text{OH}) \cdot 5\text{H}_2\text{O}$	Pale yellow-brown
Amarantite	$\text{FeSO}_4(\text{OH}) \cdot 3\text{H}_2\text{O}$	Orange red-brown red
Butlerite	$\text{FeSO}_4(\text{OH}) \cdot 2\text{H}_2\text{O}$	Deep orange
"Glockerite" (lepidocrocite?)	$\text{Fe}_4\text{SO}_4(\text{OH})_{10} \cdot n\text{H}_2\text{O}$	Brown-ochre yellow
Roemerite	$\text{Fe}^{2+}\text{Fe}_2^{3+}(\text{SO}_4)_4 \cdot 14\text{H}_2\text{O}$	Rust brown-yellow
Jarosite (alunite)	$\text{KFe}_3(\text{SO}_4)_2(\text{OH})_6$	Yellow-dark brown
Sideronatrite	$\text{K}(\text{Al}, \text{Fe})_3(\text{SO}_4)_2(\text{OH})_6$	
Metavoltine	$\text{Na}_2\text{Fe}(\text{SO}_4)_2(\text{OH}) \cdot 3\text{H}_2\text{O}$	Pale orange-yellow brown
Copiapite	$(\text{K}, \text{Na})_5\text{Fe}_3(\text{SO}_4)_6(\text{OH})_2 \cdot 8\text{H}_2\text{O}$	Yellowish brown-orange brown
Botryogen	$(\text{Fe}^{2+}, \text{Mg}, \text{Ca})\text{Fe}_4^{3+}(\text{SO}_4)_6(\text{OH})_2 \cdot 20\text{H}_2\text{O}$	Yellow-orange brown
Pyroaurite and sjogrenite	$\text{MgFe}(\text{SO}_4)_2(\text{OH}) \cdot 7\text{H}_2\text{O}$	Red-orange
Brugnatellite	$\text{Mg}_6\text{FeCO}_3(\text{OH})_{16} \cdot 4\text{H}_2\text{O}$	Yellow-brown
Rockbridgeite	$\text{Mg}_6\text{FeCO}_3(\text{OH})_{13} \cdot 4\text{H}_2\text{O}$	Orange-brown
Dufrenite	$(\text{Fe}^{2+}, \text{Mn})\text{Fe}_2^{3+}(\text{PO}_4)_3(\text{OH})_5$	Reddish brown
	$\text{Fe}^{2+}\text{Fe}_4^{3+}(\text{PO}_4)_3(\text{OH})_5 \cdot 2\text{H}_2\text{O}$	Olive brown-red brown

to their concentrations in parent basaltic rocks. These correlations suggest that Beaver Bay-type iron-rich basaltic rocks and derivative hisingerite alteration products might be analogues of martian materials.

A precursor to whole-rock alteration of iron-rich basalts is iddingsite, which is specific to olivine and frequently occurs as a reddish-brown pseudomorph after olivine in basic lavas²¹. X-ray diffraction measurements indicate that iddingsite consists of a mixture of more crystalline phases, including phyllosilicates (montmorillonite and chlorite), and an iron oxide assumed to be goethite, commonly quartz and calcite, and rarely talc and mica²². It has been suggested from textural evidence²³ that iddingsite forms from olivine phenocrysts in volatile-rich fluids before the magma has consolidated, but a more recent interpretation²⁴ is that alteration of olivine occurs at lower temperatures (<140 °C), producing an oriented intergrowth of goethite and phyllosilicates. The process of iddingsitization is not a simple process of oxidation and hydration, but results from leaching of Mg, addition of ferric iron and water, and constancy of Si content²⁵. These element fractionations are comparable to those proposed during the formation of martian regolith^{4,6,7} and again the average chemical composition of iddingsite^{21,22} shown in Table 1 (column 7) bears resemblances to that of martian fines as deduced from the Viking XRF analyses⁷, as well as that of the 'unknown' alteration products occurring in nakhlite meteorites²⁶ believed to have originated from Mars¹³.

The susceptibility of olivine to iddingsitization has important implications for remote-sensing observations of the surface of Mars, as it may lead to this mineral not being detected in reflectance spectral profiles^{27,28}. Unaltered Fe^{2+} -bearing olivines produce a distinctive crystal field transition near 1.05 μm (ref. 29). However, earth-based infrared spectra of dark areas of Mars²⁷, which are probably interfered with least by wind-blown oxidized duricrust, show absorption bands at 0.99 μm (Fe-rich augite) and 0.88 μm (Fe-rich orthopyroxene and/or Fe^{3+} -bearing phase). These spectra were interpreted as originating from a pyroxene-rich mafic rock (for example, shergottite or nakhlite) stained by ferric oxide^{4,13,27}. However, oxidization of Fe^{2+} to Fe^{3+} when olivine alters to iddingsite would be manifested by the absorption band at 0.88 μm . Whole-rock alteration to hisingerite would further eliminate Fe^{2+} -pyroxene bands, which are absent from remote-sensed spectra of martian bright areas^{28,30}.

The presence of minor amounts of sulphate, phosphate and carbonate minerals in martian regolith has been inferred from the Viking XRF measurements^{6,7}, suggesting that other Fe^{3+} -bearing minerals may also occur on the surface of Mars. Iron sulphides have been proposed as minor components of the mantle of Mars¹, and troilite and pyrrhotite have been reported in SNG meteorites¹⁴. Surface weathering of iron sulphides on

Earth, including pyrrhotite and pyrite, leads to the formation of a large variety of hydrated ferric sulphate minerals, examples of which are listed in Table 2. Also included in Table 2 are pyroaurite and brugnatellite, which are associated with weathered ultramafic rocks. Many Fe^{3+} sulphates occur in arid regions (for example, jarosite, botryogen, metavoltine): some are partially soluble in water or hydrolysed to orange-brown residues (for example, fibroferrite, botryogen); others are dehydrated at ambient temperatures (for example, hohmannite, copiapite). The colours and low-temperature dehydration products of several of the terrestrial Fe^{3+} sulphate and related minerals listed in Table 2 resemble hues and surface weathering processes described as occurring on the surface of Mars, suggesting that some of the minerals might be components of the martian duricrust.

The strong compositional and paragenetic similarities between martian regolith materials and terrestrial hisingerite and iddingsite are consistent with a mineral assemblage on the surface of Mars consisting of ferric-bearing phyllosilicate, sulphate and oxyhydroxide phases. These phases could have been derived from post-magmatic deuteric alteration of volatile-rich iron-rich basaltic rocks containing fayalitic olivine, Fe pyroxenes and accessory pyrrhotite.

This research was supported by NASA grant NSG 7604.

Received 12 September 1985; accepted 3 January 1986.

- McGetchin, T. R. & Smyth, J. R. *Icarus* **34**, 512-536 (1978).
- Morgan, J. W. & Anders, E. *Geochim. cosmochim. Acta* **43**, 1601-1610 (1979).
- Arvidson, R. E., Goettel, K. A. & Hohenberg, C. M. *Rev. Geophys. Space Phys.* **18**, 565-603 (1980).
- Baird, A. K. & Clark, B. C. *Icarus* **45**, 113-123 (1981).
- Siever, R. & Woodford, N. *Geochim. cosmochim. Acta* **43**, 717-724 (1979).
- Toulmin, P. III *et al. J. geophys. Res.* **82**, 4625-4634 (1977).
- Clark, B. C. *et al. J. geophys. Res.* **87**(B), 10059-10067 (1982).
- Clark, B. C. & van Hart, D. C. *Icarus* **45**, 370-378 (1981).
- Banin, A., Margulies, L. & Chen, Y. *J. geophys. Res.* **90**(C), 771-774 (1985).
- Gooding, J. L. & Keil, K. *Geophys. Res. Lett.* **5**, 727-730 (1978).
- Berkley, J. L. & Drake, M. J. *Icarus* **45**, 231-249 (1981).
- Allen, C. C., Gooding, J. L. & Keil, K. *J. geophys. Res.* **87**(B), 10083-10101 (1982).
- Wood, C. A. & Ashwal, L. D. *Proc. lunar planet. Sci.* **12**, 1359-1375 (1981).
- Smith, J. V., Steele, I. M. & Leitch, C. A. *J. geophys. Res.* **88**(B) 229-236 (1983).
- McSweeney, H. Y. Jr. & Stolper, E. *Scient. Am.* **242**(6), 54-63 (1980).
- Muir, I. D. *Mineralog. Mag.* **30**, 376-388 (1954).
- Grout, F. F. & Schwartz, G. M. *Bull. Minn. geol. Surv.* **28**, 119 (1939).
- Whelan, J. A. & Goldich, S. S. *Am. Miner.* **46**, 1412-1423 (1961).
- Inglis, O. G. & Willoughby, D. R. *Soil Sci.* **104**, 383-385 (1967).
- Mackenzie, K. J. D. & Berezowski, R. M. *Thermochim. Acta* **41**, 335-355 (1980).
- Deer, W. A., Howie, R. A. & Zussman, J. *Rock-Forming Minerals* Vol. 1A, 170-174 (Longman, Harlow, 1982).
- Wilshire, H. G. *Am. Miner.* **43**, 120-147 (1958).
- Edwards, A. B. *Am. Miner.* **23**, 277-281 (1938).
- Baker, I. & Haggerty, S. E. *Contr. Miner. Petrol.* **16**, 258-273 (1967).
- Gay, P. & LeMaitre, R. W. *Am. Miner.* **46**, 92-111 (1961).
- Bunch, T. E. & Reid, A. M. *Meteoritics* **10**, 303-315 (1975).
- Singer, R. B. *J. geophys. Res.* **86**, 7967-7982 (1981).
- Singer, R. B. *J. geophys. Res.* **87**(B), 3031-3042 (1982).
- Burns, R. G. *Am. Miner.* **55**, 1608-1632 (1970).
- Sherman, D. M., Burns, R. G. & Burns, V. M. *J. geophys. Res.* **87**(B), 10169-10180 (1982).

Botryococcane in a new class of Australian non-marine crude oils

D. M. McKirdy*, R. E. Cox*, J. K. Volkman† & V. J. Howell‡

* Australian Mineral Development Laboratories, Flemington Street, Frewville, South Australia 5063, Australia

† CSIRO Division of Oceanography, GPO Box 1538, Hobart, Tasmania 7001, Australia

‡ 52 Lypiatt View, Bussage, Stroud, Gloucestershire GL6 8DA, UK

There are few documented examples of source-specific hydrocarbons in non-marine petroleum, and most of these are related to land-plant precursors¹⁻³. One such biological marker of algal origin is botryococcane ($C_{34}H_{70}$), a geologically rare acyclic alkane derived from the freshwater green alga *Botryococcus braunii*^{4,5}, first discovered in several Sumatran crude oils^{6,7} and here identified in a suite of coastal bitumens from 10 different stranding sites between Kingston, South Australia and Portland, Victoria (Table 1, Figs 1, 2). Biological marker and carbon isotope analyses of these bitumens provide evidence for a new class of Australian non-marine crude oils. The bitumens are inspissated waxy residues of paraffinic and paraffinic-naphthenic crude oils that emanate from submarine seeps in the western Otway Basin⁸⁻¹², and are characterized by high concentrations of botryococcane. Unlike other high-wax Australian oils^{13,14}, the coastal bitumens contain up to 2.6% sulphur and seem to be primarily of algal origin. Their geochemistry can be used to predict the occurrence of extensive organic-rich source beds that were deposited in deep meromictic lakes along Australia's southern continental margin.

Weathered bitumen and fresh waxy crude oil are regularly washed ashore along the coastline described here⁸⁻¹⁰. The bitumens (specific gravity 5-27° API) range from paraffinic to aromatic-intermediate in bulk composition¹² and may be divided into four families on the basis of carbon isotope signature and sulphur content (Fig. 3). Three of the bitumen families are weathered, high-wax crudes that have undergone slight to moderate biodegradation (Fig. 2a). The fourth family comprises non-waxy asphaltite. These coastal bitumens can be linked to offshore oil seeps on the basis of drift bottle experiments⁹, the weathering characteristics of the bitumens¹⁰, the highly faulted nature of the continental margin in this area^{15,16}, and the fact that the quantity of stranded bitumen temporarily increases following local earthquake activity⁸.

The existence of adjacent submarine oil seepage was inferred as long ago as 1915 (refs 17, 18), when several unsuccessful petroleum exploration wells were drilled near Kingston and Robe (Fig. 1). Although subsequent onshore and offshore drilling has yet to locate commercial hydrocarbons, the coastal bitumens provide proof that oil has been generated in the subsurface of the western Otway Basin. Recent sea-borne gas 'sniffer' and sidescan sonar surveys have confirmed the existence of hydrocarbon gas plumes in the area¹¹.

Botryococcane has been identified in all the bitumens analysed by gas chromatography and mass spectrometry (Table 1, Fig. 2). *B. braunii*, the only synthesizer of botryococcane^{4,19,20} occurs as two physiologically distinct clonal races^{20,21}, both rich in hydrocarbons²²⁻²⁶. Cultivated strains of race A contain up to 61% (dry weight) of odd-carbon-numbered C_{23} - C_{31} *n*-alkadienes and trienes, whereas 24-76% of the dry cell weight of race B consists of polymethylated triterpenes of the botryococcane type (C_nH_{2n-10} ; $n = 30-37$). The two races may coexist in the same algal bloom^{19,21}. Thus the same alga is capable of contributing unsaturated hydrocarbons that are potential sources of botryococcane and long-chain *n*-alkanes to freshwater (lacustrine) sediments.

The least-weathered waxy bitumens have a bimodal *n*-alkane profile with maxima at *n*- C_{17-19} and *n*- C_{27-32} (ref. 12). The

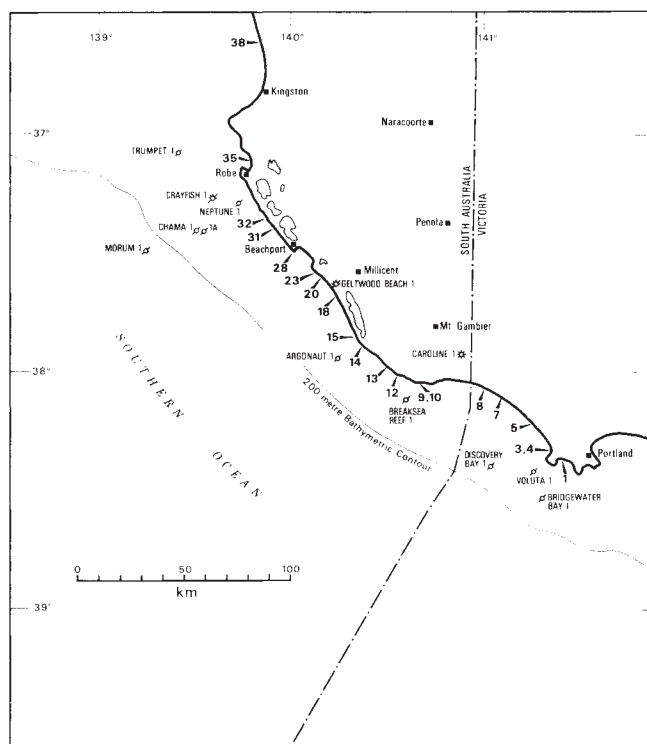


Fig. 1 Well localities and coastal bitumen stranding sites, western Otway Basin, South Australia and Victoria. Bold numbers, coastal bitumen sample locality; ○, dry, abandoned wells; ⊗, abandoned well with show of gas; ⊕, CO_2 -producing well.

lower-molecular-weight hydrocarbons are typical catagenetic products of algal lipids. Although C_{23+} *n*-alkanes in mature crude oils are usually considered to be derived from the long-chain *n*-alkanes and wax esters of higher plants, this is not true of the present crudes. Probable precursors of their waxy *n*-alkanes include the C_{23} - C_{33} *n*-alkenes found in *B. braunii* (race A) and other microscopic algae²¹⁻²³. These alkenes display a marked predominance of odd carbon-chain-lengths, which would be expected to disappear during catagenesis. Certain bacteria are potential sources of C_{24} - C_{35} *n*-alkanes with no odd/even predominance^{27,28}. Microbial activity in sediments also appears capable of rapidly degrading primary algal and/or higher plant *n*-alkanes and resynthesizing them without their original predominance of odd-carbon-numbered homologues²⁹. An additional (or alternative) source of the *n*-alkanes in the bitumens is a resistant biopolymer present in the outer cell walls of *B. braunii* (race A)³⁰. This aliphatic polymer comprises ~9% of the dry weight of the alga. It yields *n*-alkanes up to C_{31} on pyrolysis and appears to be a major building block for *Botryococcus*-derived kerogens³¹.

Carbon isotope data (Fig. 3) concur with the evidence from biomarkers in demonstrating the algal origin of the coastal bitumens. According to a recent study of 339 crude oils³², differences in source affinity (land plant or terrigenous versus marine) manifest themselves in the carbon isotope relationship between the C_{15+} saturated and aromatic hydrocarbon fractions. Sofer³² defined a canonical variable ($CV = -2.53 \delta^{13}C_{sat.} + 2.22 \delta^{13}C_{arom.} - 11.65$) which may be used to estimate the relative inputs of land-plant and algal organic matter to the source material of a particular oil. CV values >0.47 indicate a mainly terrigenous source. Otway Basin bitumens have CV values <0.47 (range -1.02 to -3.34) and therefore plot above the oblique line in Fig. 3, consistent with their predominantly algal origin. However, the precursor algae of these bitumens have a freshwater habitat. Consequently, we prefer to substitute aquatic (or

algal/bacterial) for Sofer's term 'marine' when describing non-terrestrial organic matter. Moreover, it is clear from our results that 'waxy' should not be equated with 'terrestrial'; in certain cases, waxy hydrocarbons may be entirely of algal origin.

The low C_{29}/C_{27} sterane ratios (<1.5) of the coastal bitumens preclude their derivation from vascular plants³³. On the other hand, their sterane patterns (Table 1) bear very little resemblance to the sterol distribution ($C_{27}:C_{28}:C_{29}=8:50:42$) of an extant strain of *B. braunii* analysed by one of us (J.K.V.). However, sterols are very minor constituents of *B. braunii* (J.K.V., unpublished results), so that they are unlikely to contribute significantly to the total sterane content of most sediments in which remains of the alga are preserved.

The family 1 bitumens are characterized by very high concentrations of both C_{34} botryococcane (the main component of the branched/cyclic alkane fraction, Fig. 2) and C_{30} 4-methylsteranes (Table 1). Although the side-chain structure of the 4-methylsteranes is yet to be established (either 23,24-dimethyl or 24-ethyl substitution is possible), their most likely biological precursors are 4-methylsteroids in freshwater dinoflagellates³⁴. Lesser concentrations of these source-specific hydrocarbons occur in the bitumens of families 2–4. High 4-methylsterane/sterane ratios have been reported in Tertiary lacustrine oils and shales from the Shengli field, China³⁵. Of particular interest is the co-occurrence of 4-methylsteranes and C_{31} and C_{33} botryococcane in the Maoming oil shale, another lacustrine deposit³⁶.

The primary *Botryococcus*, dinoflagellate and possibly other algal inputs to the lacustrine source rocks of the coastal bitumens were probably reworked to varying degrees by heterotrophic bacteria during deposition and early diagenesis. Major contributions of bacterial lipids to the source material of the bitumens are evident in their high hopane/sterane ratios (C_{30} hopane/ C_{29} sterane ≥ 4 : Table 1), and in the presence of significant concentrations of C_{21} – C_{31} regular (head-to-tail) and C_{37} – C_{40} irregular (head-to-head) isoprenoid alkanes (Fig. 2b). These long-chain isoprenoids are biological markers of archaeobacteria^{37–39} which, in the case of the coastal bitumens, were most probably methanogenic bacteria.

The bitumens in families 2–4 have moderately high sulphur contents (1–3%) and uniformly low ratios of the C_{19} and C_{20}

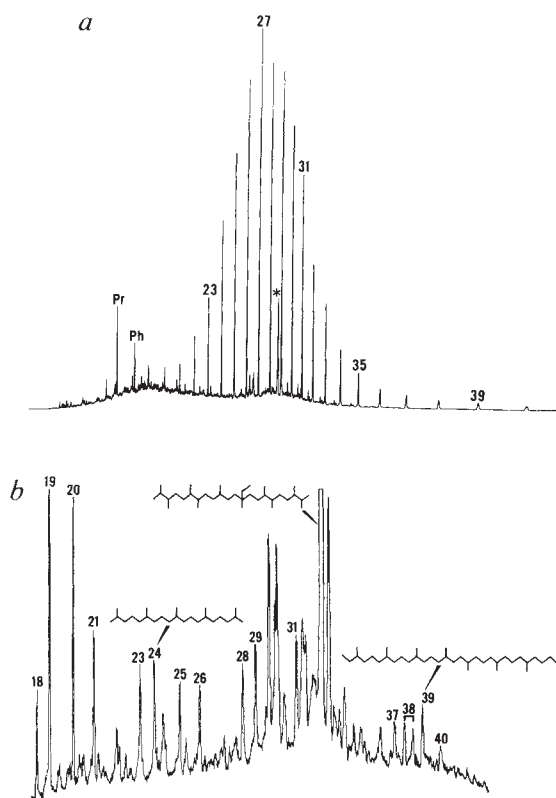


Fig. 2 Family 1 coastal bitumen from locality 38 (Fig. 1) a, Gas chromatogram of total alkanes. Pr, pristane; Ph, phytane; *, botryococcane; numbers, chain length of *n*-alkanes. GC conditions: 25 m $\times$ 0.3 mm fused silica WCOT column (SGE QC3/BP1) programmed from 80 to 290 °C at 5 °C min⁻¹ and then isothermal at 290 °C until all peaks eluted. b, Mass fragmentogram (m/z 113) of branched/cyclic alkanes; 18–31, carbon numbers of regular (head-to-tail) isoprenoids; 37–40, carbon numbers of irregular (head-to-head) isoprenoids; major peak (off scale), botryococcane. Isoprenoids and botryococcane were identified by comparing their mass spectra and GC retention times with published data^{4,37,38,41}.

Table 1 Biomarker parameters of source and maturity in selected coastal bitumens, western Otway Basin

Locality (Fig. 1)	Steranes				Triterpanes			Acyclic alkanes				
	$C_{27}:C_{28}:C_{29}$	$\frac{C_{29}}{C_{27}} \alpha\alpha$ 20R	$C_{29} \alpha\alpha$ $\frac{20S}{20R}$	C_{29} diasterane sterane	$C_{27} \frac{T_m}{T_s}$	$C_{32} \frac{22S}{22R}$	C_{30} moretane hopane	pristane phytane	botryococcane <i>n</i> -C ₂₉	botryococcane index ⁷	Hopane Sterane	4-Me sterane Sterane
Family 1												
1	74:4:22	0.34	0.89	0.93	0.40	1.3	0.10	2.2	0.27	740	53	1.8
38	78:2:20	0.26	0.75	1.3	0.41	1.3	0.09	2.0	0.45	560	57	2.6
Family 2												
31	44:6:50	1.1	0.96	0.42	0.35	1.6	0.09	0.49	—	10	4.1	0.29
Family 3												
3	39:17:44	1.1	0.83	0.27	1.0	1.5	0.08	1.0	0.06	170	6.7	0.29
4	38:18:44	1.2	0.85	0.29	1.1	1.5	0.07	1.1	—	140	7.1	0.27
8	37:18:45	1.2	0.89	0.29	1.1	1.5	0.08	0.88	—	140	5.8	0.33
10	36:20:44	1.2	0.85	0.17	1.1	1.5	0.08	1.0	0.08	110	4.8	0.31
13	36:17:47	1.3	0.78	0.22	1.1	1.5	0.08	0.86	0.10	120	5.6	0.31
18	36:19:45	1.2	0.92	0.22	1.1	1.6	0.09	1.0	—	100	5.7	0.35
Family 4												
32	40:23:37	0.92	0.79	0.30	1.4	1.4	0.12	1.1	—	3	3.5	0.19
Parameter*	1	2	3	4	5	6	7	8	9	10	11	12

* Ratios calculated from peak areas in mass fragmentograms of branched/cyclic alkanes and capillary gas chromatograms of total alkanes as follows: (1) $C_{27}:C_{28}:C_{29}$ 5 α (H)14 α (H)17 α (H) 20R-steranes ($m/z=217$). (2) C_{29} 5 α (H)14 α (H)17 α (H) 20R-sterane/ C_{27} 5 α (H)14 α (H)17 α (H) 20R-sterane ($m/z=217$). (3) C_{29} 5 α (H)14 α (H)17 α (H) 20S-sterane/ C_{29} 5 α (H)14 α (H)17 α (H) 20R-sterane ($m/z=217$). (4) C_{29} 13 β (H)17 α (H) 20R+20S-diasteranes/ C_{29} 5 α (H)14 α (H)17 α (H) 20R-steranes ($m/z=217$, 259). (5) C_{27} 17 α (H)-22,29,30-trisnorhopane(T_m)/ C_{27} 18 α (H)-22,29,30-trisnorhopane(T_g) ($m/z=191$). (6) C_{32} 17 α (H)21 β (H) 22S-homohopane/ C_{32} 17 α (H)21 β (H) 22R-homohopane ($m/z=191$). (7) C_{30} 17 β (H)21 α (H)-moretane/ C_{30} 17 α (H)21 β (H)-hopane ($m/z=191$). (8) C_{19} isoprenoid alkane/ C_{20} isoprenoid alkane (GC). (9) botryococcane/*n*-nonacosane (GC). (10) botryococcane $\times$ 100/ C_{37} – C_{40} head-to-head isoprenoid alkanes ($m/z=183$). (11) C_{30} 17 α (H)21 β (H)-hopane/ C_{29} 5 α (H)14 α (H)17 α (H) 20R+20S-4-methylsteranes/ C_{27} + C_{29} 5 α (H)14 α (H)17 α (H) 20R-steranes ($m/z=217$, 231). Branched/cyclic alkanes analysed in multiple-ion detection mode (electron energy 29 eV, mass range 50–450 AMU, scan time 1 s) using a Finnigan 3200 quadrupole gas chromatograph-mass spectrometer (GC-MS) interfaced to an INCOS 2015 (Nova 4) data system. GC conditions: 25 m $\times$ 0.32 mm fused silica WCOT OV-1 cross bond (Mega) column; temperature programme 100–150 °C at 10 °C min⁻¹, 150–300 °C at 6 °C min⁻¹; on column injection.

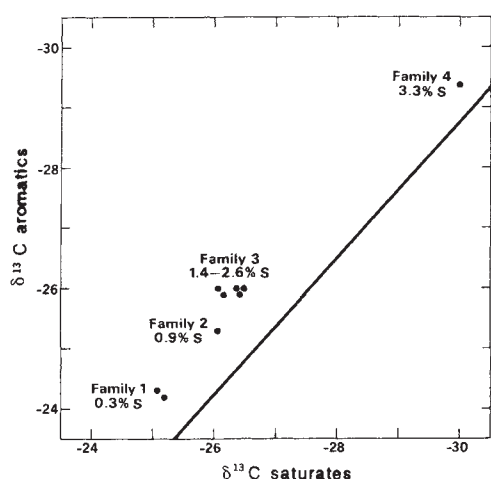


Fig. 3 Stable carbon isotope composition and sulphur content of coastal bitumens, western Otway Basin. The $\delta^{13}\text{C}$ (‰PDB) values plotted were measured on C_{12+} saturated hydrocarbon (alkane) and aromatic hydrocarbon fractions isolated from deasphalted bitumens by liquid chromatography on activated silica gel and alumina. The oblique line separates land-plant-derived (below line) from algal-derived oils (above line)³².

isoprenoid alkanes, pristane and phytane (0.5–1). This combination of geochemical features^{1,13} reflects deposition of their algal source material under stable anoxic conditions. Somewhat less reducing conditions characterized the source-bed environment of the family 1 bitumens (pristane/phytane = 2; S = 0.3%). The progressive increase in sulphur content from paraffinic family 1 bitumens to asphaltic family 4 bitumens (Fig. 3) implies increasing availability of H_2S (from bacterial sulphate reduction) for chemical or bacterial oxidation to elemental sulphur and eventual incorporation into their source-rock kerogens³⁹. This may have required development of a saline (sulphate-rich, iron-poor) hypolimnion, particularly for those sediments which became the source of the family 4 bitumen.

Elsewhere in the Otway Basin (for example, Port Campbell-4 and Lindon-1), and in other basins throughout Australia, high-wax oils have low sulphur contents (<0.5%), generally high pristane/phytane ratios (4–10) and are found in coal-bearing or deltaic sedimentary sequences^{13,14}. Hence, McKirdy and Horvath, on first observing the low pristane/phytane ratios of the waxy Otway Basin coastal bitumens, concluded erroneously that they were examples of immature crude oils derived from terrigenous (land plant) organic matter¹⁰. We show here that they are, in fact, mature lacustrine oils primarily of algal origin. The C_{29} sterane $\alpha\alpha\alpha$ 20S/20R ratios of the coastal bitumens (Table 1) are consistent with migration from their respective source rocks at vitrinite reflectance levels in the range 0.65–1% (ref. 2).

The presence of C_{34} botryococcane in the coastal bitumens suggests that the parent oils of the bitumens were generated from lacustrine source beds. This represents the first published identification of botryococcane in an Australian crude oil. Preservation of the remains of *Botryococcus* blooms in sediments deposited under anoxic to micro-oxic conditions in deep stratified lakes accounts for both the waxy character of the three main bitumen types and their botryococcane content.

Regional lithofacies studies of the early Cretaceous Otway Group (J. G. G. Morton and G. R. McClung, personal communications) suggest that the lacustrine source rocks from which the coastal bitumens originated are located deep in the western offshore Otway Basin. Such sediments could have accumulated in large meromictic lakes which formed in rift valleys associated with the continental breakup of Australia and Antarctica^{15,16}. Major oil accumulations have been discovered in similar Cretaceous lacustrine source-reservoir sequences in rift basins

along the South Atlantic margin in Brazil (Reconcavo Basin), Gabon (Gabon Basin) and Angola (Cabinda Basin)⁴⁰.

We thank Ultramar Australia Inc., Elf Aquitaine Petroleum Australia Pty Ltd and their respective partners in exploration permits EPP 18 and PEL 8, for permission to publish. The bitumen samples were collected by Mr D. R. Vinall during August 1983.

Received 9 September 1985; accepted 7 January 1986.

1. Tissot, B. P. & Welte, D. H. *Petroleum Formation and Occurrence* (Springer, Berlin, 1984).
2. Mackenzie, A. S. *Adv. Petrol. Geochim.* **1**, 115–214 (1984).
3. Moldowan, J. M., Seifert, W. K. & Gallegos, E. J. *Bull. Am. Ass. Petrol. Geol.* **69**, 1255–1268 (1985).
4. Maxwell, J. R., Douglas, A. G., Eglinton, G. & McCormick, A. *Phytochemistry* **7**, 2157–2171 (1968).
5. Cox, R. E., Burlingame, A. L., Wilson, D. M., Eglinton, G. & Maxwell, J. R. *JCS chem. Commun.*, 284–285 (1973).
6. Moldowan, J. M. & Seifert, W. K. *JCS chem. Commun.*, 912–914 (1980).
7. Seifert, W. K. & Moldowan, J. M. *Geochim. cosmochim. Acta* **45**, 783–794 (1981).
8. Sprigg, R. C. & Woolley, J. B. *Trans. R. Soc. S. Aust.* **86**, 67–103 (1963).
9. Sprigg, R. C. *Aust. Petrol. Explor. Ass. J.* **4**, 53–63 (1964).
10. McKirdy, D. M. & Horvath, Z. *Aust. Petrol. Explor. Ass. J.* **16**(1), 123–135 (1976).
11. Sprigg, R. C. in *The History of the Oil Industry in South Australia*, C1–C119 (Aust. Inst. Petrol., Adelaide, 1983).
12. McKirdy, D. M. in *Orway 85, Earth Resources of the Orway Basin*, 34–49 (Geol. Soc. Aust., Adelaide, 1985).
13. Powell, T. G. & McKirdy, D. M. *Bull. Am. Ass. Petrol. Geol.* **59**, 1176–1197 (1975).
14. Thomas, B. M. *Aust. Petrol. Explor. Ass. J.* **22**(1), 164–178 (1982).
15. Bouef, M. G. & Doust, H. *Aust. Petrol. Explor. Ass. J.* **15**(1), 33–43 (1975).
16. Deighton, I., Falvey, D. A. & Taylor, D. J. *Aust. Petrol. Explor. Ass. J.* **16**(1), 25–36 (1976).
17. Wade, A. *Geol. Surv. S. Aust. Bull.* **4**, 1–54 (1915).
18. Colyer, F. *Petrol. Gaz.* **18**, 58–63 (1974).
19. Wake, L. V. & Hillen, L. W. *Aust. J. mar. Freshwat. Res.* **32**, 353–367 (1981).
20. Metzger, P., Casadevall, E., Pouet, M. J. & Pouet, Y. *Phytochemistry* **24**, 2995–3002 (1985).
21. Metzger, P., Berkalo, C., Casadevall, E. & Coute, A. *Phytochemistry* **24**, 2305–2312 (1985).
22. Brown, A. C., Knights, P. A. & Conway, E. *Phytochemistry* **8**, 543–547 (1969).
23. Gelpi, E., Schneider, H., Mann, J. & Oro, J. *Phytochemistry* **9**, 603–612 (1970).
24. Knights, B. A., Brown, A. C., Conway, E. & Middleditch, B. S. *Phytochemistry* **9**, 1317–1324 (1970).
25. Galbraith, M. N., Hillen, L. W. & Wake, L. V. *Phytochemistry* **22**, 1441–1443 (1983).
26. Metzger, P. & Casadevall, E. *Tetrahedron Lett.* **24**, 4013–4016 (1983).
27. Davis, J. B. *Chem. Geol.* **3**, 155–160 (1968).
28. Albro, P. W. & Dittmer, J. C. *Lipids* **5**, 320–325 (1970).
29. Johnson, R. W. & Calder, J. A. *Geochim. cosmochim. Acta* **37**, 1943–1955 (1973).
30. Berkalo, C. *et al. Phytochemistry* **22**, 389–397 (1983).
31. Largeau, C., Casadevall, E., Kadouri, A. & Metzger, P. *Org. Geochem.* **6**, 327–332 (1984).
32. Sofer, Z. *Bull. Am. Ass. Petrol. Geol.* **68**, 31–49 (1984).
33. Meinschein, W. G. & Huang, W. Y. in *Origin and Chemistry of Petroleum* (eds Atkinson, G. & Zuckerman, J. J.) 33–55 (Pergamon, Oxford, 1981).
34. Robinson, M., Eglinton, G., Brassell, S. C. & Cranwell, P. A. *Nature* **308**, 439–442 (1984).
35. Shi Ji-Yang *et al. Chem. Geol.* **35**, 1–31.
36. Brassell, S. C., Eglinton, G. & Fu Jia Mo *Org. Geochem.* (submitted).
37. Moldowan, J. M. & Seifert, W. K. *Science* **204**, 169–171 (1979).
38. Albaiges, J. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 19–28 (Pergamon, Oxford, 1980).
39. Lijmbach, G. W. M. *Proc. 9th Wild. Petrol. Congr.* **2**, 357–369 (1975).
40. Robbins, E. I. *Tectonophysics* **94**, 633–658 (1983).
41. Holzer, G., Oro, J. & Tornabene, T. G. *J. Chromat.* **186**, 795–809 (1979).

Age of Pleistocene faunas from Bacon Hole, Wales

C. B. Stringer, A. P. Currant, H. P. Schwarcz* & S. N. Collcutt†

Department of Palaeontology, British Museum (Natural History), Cromwell Rd, London SW7 5BD, UK

* Department of Geology, McMaster University, Hamilton, Ontario, Canada L8S 4M1

† D. Baden-Powell Centre for Quaternary Research, Banbury Rd, Oxford OX2 6PN, UK

At present it is difficult to match British early Upper Pleistocene vertebrate faunas with the growing evidence for complex climatic change at this time derived from the oceanic oxygen isotope record and from long pollen cores. However, we report here new analyses of dated faunal assemblages from Bacon Hole, Wales, which indicate that the Ipswichian interglacial fauna dated at ~125 kyr BP was followed by a Devensian interstadial fauna dated at ~81 kyr BP, which still supported some supposedly typical interglacial mammals. Combined with evidence of an intervening 'cold'

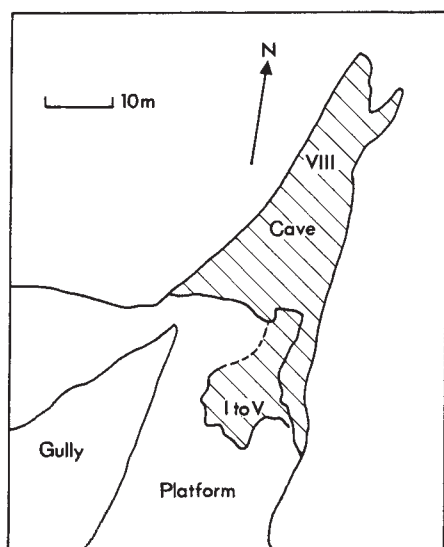


Fig. 1 Simplified plan of Bacon Hole. Shaded area shows the extent of the surviving cave sediments.

fauna from Stump Cross Cave dated at more than 83 kyr BP, these indications of complexity in faunal change at the Ipswichian-Devensian transition are more consistent with the oxygen isotope and pollen core data than traditional views.

In Britain, the period of time covered by the late Pleistocene is less well represented by datable deposits than the equivalent period in continental Europe, where continuous cores have produced long palynological records showing close correspondence between changes in vegetation patterns and the climatic inferences based on the oceanic oxygen isotope record¹⁻³. From about 130 kyr BP to 73 kyr BP we find an interglacial stage (5e of the oxygen isotope record = Eemian of continental Europe), followed by two less fully expressed temperate stages (probably equivalent to 5c and 5a), with intervening cold stages (probably equivalent to 5d and 5b). Thereafter, there is evidence for an extended cold stage¹⁻³.

Oxygen isotope and biotic data show that during stage 5e, centred around 122 kyr BP, North Atlantic temperatures were similar to or greater than those of the Holocene⁴. While the oxygen isotope record can determine relative global ice volume, and hence relative sea level, it does not indicate actual sea level; this information must be sought from sources such as dated coral reefs. A sea level of ~5 m above present levels, dated at ~125 kyr BP, has been determined on the relatively stable carbonate platform of Bermuda. Stage 5e was the last period when fully interglacial conditions of ocean temperature, ice volume, sea level and continental vegetation patterns prevailed prior to the Holocene^{4,5}.

In Britain, the climatic events corresponding to oxygen isotope stages 5d to 5a have not been formally recognized, even though the Devensian cold stage is known to include early interstadials of uncertain date and status^{6,7}. Existing models of mammalian faunal history during the Ipswichian and Early Devensian adopt a composite palynological framework in which the Ipswichian is seen as a single climatic and biotic cycle with a smooth transition into the succeeding cold stage. In the absence of absolute dates and sites with long faunal sequences, isolated faunas have been ordered and correlated on the basis of associated palynological data⁸. In this scheme mammals such as *Hippopotamus*, *Dama*, *Palaeoloxodon* and *Dicerorhinus* characterize the climatic optimum (pollen zone Ip IIB), giving way to forms such as *Mammuthus*, *Equus* and *Coelodonta* by the end of the stage (pollen zone Ip IV). These then continue into the Devensian, supplemented by forms such as *Rangifer*, but without further major faunal changes until after the late Devensian

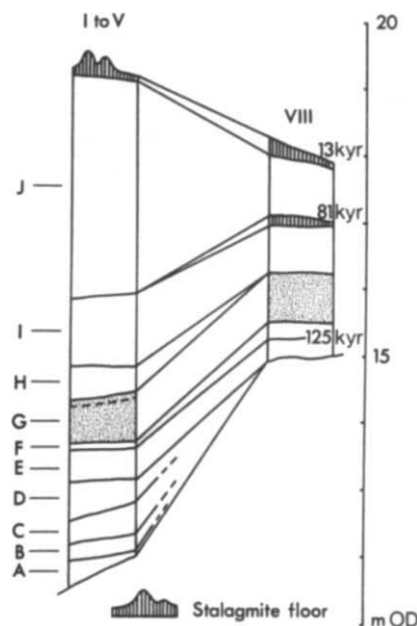


Fig. 2 Basal Pebbles (A): possibly a coarse littoral lag deposit but no longer *in situ*. Coarse Grey (B) and Coarse Orange (C) Sands: evidence of wind deposition and a fauna with *Equus ferus* and a large form of *Microtus oeconomus* suggesting relatively cold conditions⁸; land snails indicate bare rock and scree with little vegetation. Sandy Breccio-Conglomerate (D): pebbles, shingle and coarse sand with shell debris, probably a storm beach representing a sea level only marginally higher than that of today. Sandy Cave Earth (E): coarse angular heavily altered limestone clasts in a heterogeneous matrix with sandy and shelly lenses and pockets of land snails. At the base of the area VIII sequence (mean age given in kyr) it appears to have been capped by a stalagmite floor which was subsequently broken up and incorporated into the succeeding unit, the Shelly Sand (F) which comprises primarily shell debris of small littorinids. The associated vertebrate fauna is consistent with local temperate woodland. The Grey Clays, Silts and Sands (G): marked syn- and post-depositional modification by waterlogging and downslope creep. The associated bone and coprolite accumulation is largely due to *Crocota* occupation. Upper Sands (H): well-bedded medium sands with much fine shell debris, probably accumulated by wind action. Foraminifera include the last occurrence of an extinct species of *Rosalina* also known from other Eemian/Ipswichian sites in Europe. Upper Cave Earth (I): a return to an environment similar to that represented by the Sandy Cave Earth but with minimal marine influence. At one stage this deposit blocked the cave entrance, initiating stalagmite deposition (sampled as BH 1981-250; Table 1). The resultant date is taken as a good approximation to the age of the Upper Cave Earth and its fauna. Cemented Breccias (J): complex of upward-fining cyclothem forming a marked talus cone in the present entrance and capped by a massive stalagmite floor.

glacial maximum. The assumptions behind this model do not allow for a complex transition between the Ipswichian and Devensian such as might be expected from other lines of evidence. Faunas assigned to the later part of the interglacial are particularly controversial⁸⁻¹⁰ as none have been found stratigraphically superimposed on characteristic earlier Ipswichian faunas. Aveley in Essex, the one site where such a relationship is claimed, has a climatic optimum fauna unlike other Ipswichian assemblages⁶.

The distinctive Ipswichian *Hippopotamus* fauna is associated with pollen indicating an interglacial climatic optimum^{8,9,11} and dated at or before 120 kyr BP at Victoria Cave, North Yorkshire¹². Dating of the 8-m raised beach at Belle Hogue, Jersey, gives a comparable age (121 ± 12 kyr BP)¹³. Thus, the last major Pleistocene marine transgression in the British Isles and the climatic optimum of the Ipswichian can be correlated with

Table 1 Uranium series ages on stalagmite samples from Bacon Hole area VIII (west)

Sample no.	$^{234}\text{U}/^{238}\text{U}$	$^{230}\text{Th}/^{234}\text{U}$	$^{230}\text{Th}/^{232}\text{Th}$	(U) (p.p.m.)	Age in kyr	Age in kyr. (corrected)	Yields Th U	Stratigraphic significance
1978-801:01	1.498 ± 0.064	0.123 ± 0.013	15.5 ± 1.9	0.14	14.0 ± 2.0	13.0 ± 3.0	44 52	Broken block of surface stalagmite. Minimum age for Devensian fauna
:02	1.519 ± 0.040	0.159 ± 0.008	3.8 ± 0.5	0.19	18.6 ± 1.0	12.8 ± 1.7	54 58	
1981-250	1.485 ± 0.242	0.547 ± 0.087	>1,000	0.03	81 ± 18		59 39	Minimum age for 'interglacial' elements in Upper Cave Earth including partly enclosed <i>Palaeoloxodon</i> tusk
1981-212:01 (top)	1.352 ± 0.094	0.725 ± 0.048	46.7 ± 21.4	0.07	129 ± 16		56 29	All are from broken blocks of stalagmite floor incorporated in Shelly Sand. Sediment analysis suggests that the stalagmites formed on underlying Sandy Cave Earth. Dates relate to main interglacial fauna and last major Pleistocene marine transgression at the site
:02 (middle)	1.608 ± 0.172	0.762 ± 0.074	8.00 ± 1.20	0.05	136 ± 23	125 ± 26	78 25	
:03 (bottom)	1.479 ± 0.154	0.772 ± 0.078	7.6 ± 1.0	0.04	142 ± 27	129 ± 30	109 47	
1981-252:02	1.913 ± 0.178	0.704 ± 0.067	9.4 ± 2.5	0.02	116 ± 18	107 ± 21	28 38	
:01	1.343 ± 0.073	0.701 ± 0.038	31.2 ± 9.2	0.07	122 ± 11		45 38	
Mean of last five determinations (uncorrected/corrected)	1.427 ± 0.048	0.723 ± 0.025			$127 \pm \frac{9}{8}$	122 ± 9		

oxygen isotope stage 5e. Aminostratigraphy extends this correlation to the late Pleistocene 'Patella Beach' of South Wales, present at Bacon Hole and Minchin Hole¹⁴⁻¹⁶. Bacon Hole¹⁷⁻¹⁹ and Minchin Hole²⁰ are coastal caves which have produced three different lines of evidence relating to later Pleistocene events: raised beach deposits and marine influence in sediments and faunas, distinct faunas in stratigraphic succession, and stalagmite floors providing absolute dates. The sea level evidence is best represented at Minchin Hole²⁰. Bacon Hole provides a more complete faunal record for the Ipswichian and Early Devensian, as well as chronological control.

Three stalagmitic horizons were located inside Bacon Hole and samples for dating by the $^{230}\text{Th}/^{234}\text{U}$ method were collected (Table 1). Samples were dissolved in 2N HNO_3 with a tracer of ^{232}U - ^{228}Th , U and Th separated by ion exchange, plated on steel discs and analysed by α -spectrometry. Corrections for detrital ^{230}Th were based on the observed activity of ^{232}Th , assuming an initial $^{230}\text{Th}/^{232}\text{Th}$ ratio of 1.25²¹. The large uncertainties in quoted ages are due to low U concentrations in all samples. Analyses of stalagmite fragments in the Shelly Sand date a floor which, judging by adherent matrix, formed on the Sandy Cave Earth prior to deposition of the Shelly Sand. These two units and the underlying Sandy Breccio-Conglomerate show marked marine influence and close faunal similarities (Table 2). They are all correlated with oxygen isotope stage 5e. The mean age of the dated samples (weighted by size of standard error) is 127 ± 9 kyr BP with an unweighted mean (including corrected ages) of 122 ± 9 kyr BP. A single age determination on stalagmite capping the Upper Cave Earth has a large error range but is significantly younger than the stage 5e determinations for the Shelly Sand; this would be expected given three intervening units with evidence of lower sea levels and distinct faunas and associated environmental information. Attributing the stalagmites to stage 5a would be consistent with the age of 81 kyr BP but the large error precludes exact assignment. Determinations on the surface stalagmite give a Late Devensian stage 2 age, providing a minimum age for the underlying fauna and an actual date for enclosed bones of *Rangifer*.

By combining litho-, bio- and chronostratigraphic data, we can reconstruct Late Pleistocene events as recorded at Bacon Hole (Figs 1, 2; Tables 1, 2). Cave deposits apparently accumulated before stage 5e, as indicated by age determination of stalagmite capping heavily altered sediments in a cleft in the cave floor (BH 1980-stal 2, age 175 ± 19 kyr BP), but these are assumed to have been largely removed by the scouring of a pre-5e stage high sea level (perhaps represented by the remnant Basal Pebbles). After a period of cool, dry climate with relatively open vegetation (Coarse Grey and Coarse Orange Sands) there was a rise in relative sea level, temperature and humidity, with deposition of the Sandy Breccio-Conglomerate, Sandy Cave

Earth and Shelly Sand at about 125 kyr BP. These units show high levels of marine influence in the form of marine-modified sand grains and remains of molluscs, crustaceans, fish and sea birds. The faunal content and absolute dating of these units allow the first direct correlation of the British Ipswichian fauna with the high sea level event of oxygen isotope stage 5e. The Sandy Breccio-Conglomerate represents a storm beach, marginally higher than its modern equivalent, which can be correlated on altitudinal, lithostratigraphic, aminostratigraphic^{14,15,22}, chronostratigraphic and faunal grounds with the +2-m Patella Beach of Minchin Hole^{20,22}. The Sandy Breccio-Conglomerate, Sandy Cave Earth and Shelly Sand have produced evidence of a nesting population of *Calonectris diomedea* over 5° N of its present breeding range²³.

Falling sea level and temperature but increasing humidity is inferred from marked changes in lithology in the succeeding Grey Clays, Silts and Sands. *Mammuthus* is recorded for the first and only time in the sequence, while the birds of the preceding units are replaced by species such as *Anser fabalis* and *Clangula hyemalis*, now winter visitors to Britain. The Upper Sands probably represent a drier phase with the deposition of partly marine-derived wind-blown sediments. The Upper Cave Earth was certainly deposited at a time of lower sea level than the superficially similar Sandy Cave Earth and contains a comparatively restricted mammalian fauna of 'interglacial' character

Table 2 Significant fauna and inferred absolute ages for selected units at Bacon Hole

Unit	Age (kyr BP)	Mollusca	Gulo	Crocuta	Palaeoloxodon	Mammuthus	Dicerorhinus	Equus	Dama	Rangifer	M. oecornus	Apodemus	Calonectris
J: Cemented Breccias	13+		✓							✓	✓		
I: Upper Cave Earth	81+			✓	✓	✓	✓			✓	✓		
G: Grey Clays, etc.				✓	✓	✓	✓			✓	✓		
F: Shelly Sand/													
E: Sandy Cave Earth	125+	M, G			✓				✓			✓	✓
D: Sandy Breccio- Conglomerate		M, G										✓	✓
B/C: Coarse Sands		R						✓			✓		

M, Marine and littoral mollusca common; G, grassland snail fauna dominated by *Clausilia*, with some *Discus rotundatus*, *Pupilla*, *Vallonia*, *Cepaea*, cf. *Nesovittrea hammonis* and cf. *Helicella itala*; R, restricted land snail fauna dominated by *Laura cylindracea*, *Pyramidula rupestris* and *Pupilla muscorum*, with *Cochlicopa lubricella*, *Oxychilus* cf. *alliaris*, *Vitrina* sp. and *Cepaea* sp.

which probably represents an early Devensian interstadial.

If the Ipswichian *sensu stricto* is regarded as a short event corresponding to the period of high sea level and climatic optimum of oxygen isotope stage 5e (~127–115 kyr BP)^{1–5} the units post-dating the Shelly Sand probably fall into the early part of the Devensian. While the early part of the sequence supports existing concepts of faunal history up to the climatic optimum of the Ipswichian^{8,9}, subsequent faunas at Bacon Hole are novel. Most notable in this respect is the continued abundance of *Palaeoloxodon* and *Dicerorhinus* in the early Devensian (late stage 5) fauna of the Upper Cave Earth. In common with the Ipswichian mammal fauna, *Equus ferus* is not represented in the Upper Cave Earth assemblage, even though Wood^{24,25} excavated large areas of this deposit inside the cave and found abundant mammalian remains. These characteristics serve to falsify existing models of the Ipswichian/Devensian faunal transition, replacing a composite sequence with an actual one.

Evidence for complexity of faunal change in the early Devensian comes from Stump Cross Cave, North Yorkshire²⁶, where flowstone dated at 83 ± 6 kyr BP encapsulates remains of *Rangifer* and *Gulo*. This date closely follows the suggested age range of oxygen isotope stage 5b (~95–85 kyr BP) to which the cold fauna has been assigned. Remains of *Dicrostonyx* and *Microtus oeconomus* apparently mixed with more characteristically interglacial biota at Elderbush Cave, Staffordshire²⁷, may represent a similar brief cold event. The observation that a fauna with clearly cold-adapted mammals at Stump Cross Cave may have been followed by one with marked interglacial affinities as in the Upper Cave Earth at Bacon Hole, all within the early parts of the Devensian, marks a significant departure in the interpretation of Late Pleistocene faunal history in Britain. This departure should allow reassessments of other British evidence, in line with those that have already taken place elsewhere in Europe following the recognition of a greater complexity in climatic and biotic change during the Upper Pleistocene.

We thank all those who have worked with us, particularly John Evans for study of molluscs, Peter Bull for scanning electron microscopy of quartz grains²⁸, John Catt for mineralogical analyses, Jill Cook for study of polished bones, John Whitaker and colleagues for study of foraminifera, Colin Harrison for work on the avifauna and Miro Ivanovich for advice. H.P.S. acknowledges financial support from SSHRC (Canada) and analytical assistance from Nicki Cesar.

Received 29 March 1985; accepted 27 January 1986.

1. Woillard, G. M. & Mook, W. G. *Science* **215**, 159–161 (1982).
2. Beaulieu, J.-L. de & Reille, M. *Geogr. phys. Quat.* **38**, 3–9 (1984).
3. Turon, J.-L. *Nature* **309**, 673–676 (1984).
4. CLIMAP project members. *Quat. Res.* **21**, 128–224 (1984).
5. Harmon, R. S. et al. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **44**, 41–70 (1983).
6. Shotton, F. W. *IGCP Project 73/1/24 Quaternary Glaciation in the Northern Hemisphere* Rep. 9, 109–115 (Paris, 1983).
7. Shotton, F. W. *Phil. Trans. R. Soc. B* **280**, 107–118 (1977).
8. Stuart, A. J. *Pleistocene Vertebrates in the British Isles* (Longman, London, 1982).
9. Stuart, A. J. *Phil. Trans. R. Soc. B* **276**, 221–250 (1976).
10. Sutcliffe, A. J. & Kowalski, K. *Bull. Br. Mus. nat. Hist. (Geol.)* **27**(2), 1–147 (1976).
11. Gibbard, P. L. & Stuart, A. J. *Geol. Mag.* **112**, 493–501 (1975).
12. Gascoyne, M., Currant, A. P. & Lord, T. C. *Nature* **294**, 652–654 (1981).
13. Keen, D. H., Harmon, R. S. & Andrews, J. T. *Nature* **289**, 162–164 (1981).
14. Davies, K. H. *Nature* **302**, 137–139 (1983).
15. Andrews, J. T., Bowen, D. Q. & Kidson, C. *Nature* **281**, 556–558 (1979).
16. Bowen, D. Q., *Geol. J. Spec. Iss.* **7**, 228–256 (1977).
17. Allen, E. E. & Rutter, J. G. *A survey of Gower Caves with an Account of Recent Excavations* (Thomas, Swansea, 1946).
18. Stringer, C. B. *Gower* **26**, 32–37 (1975).
19. Stringer, C. B. *Gower* **28**, 34–44 (1977).
20. Sutcliffe, A. J. *Q. J. Geol.* **33**, 1–17 (1981).
21. Schwarcz, H. P. *Archaeometry* **22**, 3–24 (1980).
22. Henry, A. thesis, Univ. Wales (1984).
23. Harrison, C. J. *Trans. Br. Cave Res. Ass.* **4**, 441–442 (1977).
24. Benson, S. *Annual Report of the Swansea Literary and Scientific Society*, 10–19 (1852).
25. Murchison, C. (ed.) *Palaeontological Memoirs and Notes of the Late Hugh Falconer* (Hardwicke, London, 1868).
26. Sutcliffe, A. J. et al. *Q. J. Geol.* **43**, 9–12 (1984).
27. Bramwell, D. & Shotton, F. W. *Q. J. Geol.* **38**, 7–13 (1982).
28. Bull, P. A. Z. *Geomorph. N.F.* **36**, 217–232 (1980).

Digestive physiology is a determinant of foraging bout frequency in hummingbirds

Jared M. Diamond*, William H. Karasov*†, Duong Phan* & F. Lynn Carpenter‡

* Physiology Department, University of California Medical School, Los Angeles, California 90024, USA

‡ Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92677, USA

Hummingbirds are among the smallest endothermic vertebrates. Because they forage by energetically costly hovering, and because weight-specific basal metabolic rates increase with decreasing body size, their basal and active metabolic rates are among the highest recorded. Hummingbirds fuel these metabolic requirements mainly with highly concentrated sugar in nectar, which they extract rapidly and efficiently^{1,2} by an unknown mechanism. It is especially puzzling that, despite their high energy requirements, hummingbirds spend only ~20% of their waking hours feeding, but 75% perched and apparently doing nothing^{3,4}. Here we report the first measurement of nutrient absorption by hummingbird intestine and present a new method for measuring crop-emptying times. We find that hummingbird intestine has the highest active glucose transport rate and lowest passive glucose permeability reported for any vertebrate. Crop-emptying time may limit feeding-bout frequency and could largely account for the time spent perched.

We studied Rufous and Anna's Hummingbirds (*Selasphorus rufus*, *Calypte anna*) caught from the wild. Both species gave similar results⁵. We first made timed collections of excreta from caged birds fed 100 µl of sugar water (0.585 M sucrose) marked with either ³H-labelled polyethylene glycol and ¹⁴C-D-glucose or red dyes (Schilling FDA nos 3 and 40). The markers appeared in excreta within 15 min (Fig. 1), and mean retention time⁶ for polyethylene glycol was 49 ± 3 min (mean $\pm$ s.e.m. for 8 birds). These are relatively rapid transit times, rivalled among vertebrates only by values for phainopeplas eating berries⁷. (Times for most bird species are 1–3 h, and for most mammals and lizards 2–150 h; ref. 8.) Despite this rapid transit, intestinal extraction of ingested D-glucose, as calculated by comparing the D-glucose/polyethylene glycol ratio in sugar water with excreta, was nearly complete ($97 \pm 0.3\%$, $n = 8$) (compare with ref. 2).

To calculate the contribution of crop-emptying time to the transit time, we devised a double-isotope dilution technique for measuring crop contents at various times after a sugar-water meal. Birds were initially fed 100 µl of ¹⁴C-polyethylene glycol-labelled sugar water through a capillary tube, then they were given an additional 10 or 100 µl of ³H-polyethylene glycol-labelled sugar water after 2–20 min. After a further 5 min, which was enough time to allow uniform mixing of the two isotopes within the crop, we sampled the crop contents with a 10-µl capillary tube introduced orally. (During the experiment the birds were kept in the dark in covered cages and sat quietly.) In effect, dilution of ¹⁴C-polyethylene glycol in the sample measures crop volume at time $t = 0$, whereas dilution of ³H-polyethylene glycol measures crop volume at $t = 2$ –20 min. The apparent volume of crop contents V_t at time t was calculated from marker dilution as $(V_s C_t / C_s) - V_i$, where V_s is the analysed sample volume (10 µl), V_i is the volume of fed marker solution, C_s is the c.p.m. in the sample volume and C_t is the c.p.m. in the fed solution.

After a 100-µl meal, V_t declined to 26 ± 1 µl ($n = 12$) and remained at this value from 30 to at least 90 min, indicating that

† Present address: Department of Wildlife Ecology, University of Wisconsin, Madison, Wisconsin 53706, USA.

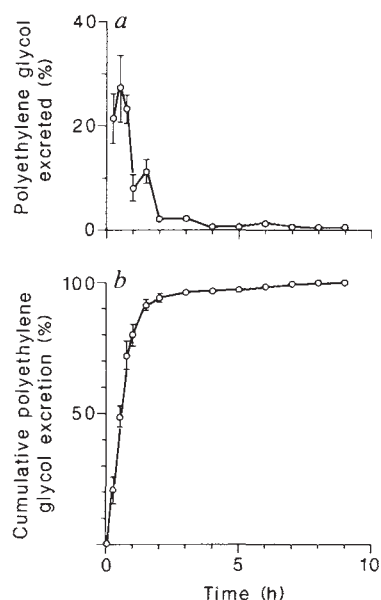


Fig. 1 Excretion of the liquid marker polyethylene glycol by Rufous Hummingbirds as a function of time after the marker was ingested. *a*, Percentage of total excreted polyethylene glycol excreted each 15 or 30 min (mean value $\pm$ s.e.m. for 4 birds). *b*, Cumulative excretion up to a given time. Note that excretion rates peak, and cumulative excretion reaches 50%, in $\sim\frac{1}{2}$ h.

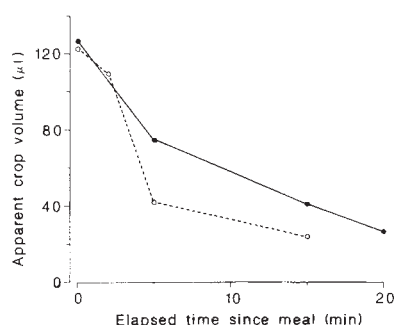


Fig. 2 Apparent volume of crop contents in two fasted Rufous Hummingbirds as a function of time elapsed after drinking 100 μ l of a sucrose solution. A crop with an apparent volume of 26 μ l is considered empty (see text). Note that the crop is emptied within $\sim$ 15–20 min.

the crop was as empty as possible. We plotted V_t against t for two birds (Fig. 2); five other birds yielded similar results. Compared with the empty volume of 26 μ l, V_t at 2–10 min was always much greater (42–110 μ l), at 15 min still significantly greater (51 ± 11 μ l, $n = 4$), and at 20 min not significantly different (33 ± 12 μ l, $n = 5$). Thus, a 100- μ l meal was cleared from the crop in 15–20 min, with a half-time (calculated from a negative exponential model⁹) of 4.1 min.

We measured intestinal absorption of D-glucose, L-proline and L-glucose *in vitro* at 37 °C, using an everted sleeve technique described previously¹⁰. (D-Glucose and L-proline transport decreased distally along the intestine, more markedly for the former than the latter.) Initially, we wondered whether, because of the very high sugar concentration in nectar meals, hummingbird intestine had evolved a high Michaelis-Menten constant, K_m (half-saturating concentration), for active D-glucose transport, or was even able to absorb glucose passively down its concentration gradient. In fact, the apparent K_m of 4.3 ± 1.1 mM ($n = 4$) was similar to that of other terrestrial vertebrates¹¹, whereas glucose passive permeability measured with L-glucose was immeasurably low, lower than in any other vertebrate studied^{11,12}. This low passive permeability of hummingbird

intestine could be an adaptation to prevent loss of blood solutes to the intestinal lumen and excreta, given the rapid and continuous fluid transit through the gut. Although the L-proline transport rate and the K_m for active glucose transport are normal in hummingbird intestine, the V_{max} of active glucose transport is higher than in any other vertebrate studied (1,570 nmol min⁻¹ cm⁻²). Thus, hummingbird intestine has adapted in two ways to a high-sugar diet: by suppressing passive permeability and by maximizing the density of glucose transport sites.

A key question in behavioural ecology is whether an animal is an energy maximizer that spends as much time foraging as possible, or a foraging-time minimizer that cannot profit from additional food intake beyond a certain level¹³. Field observations suggest that hummingbirds are time minimizers: Rufous Hummingbirds feed 14–18 times per h (ref. 4), taking <1 min per bout and spending most of their time perched seemingly inactively^{3,4,14,15}. However, our results suggest that they are using that time to digest as fast as possible and that they cannot feed again until they have cleared their crop from the previous bout. Assuming that a bird clears half of a meal before it forages again, our estimate of 4.1 min as the half-time for clearing 100 μ l means that the bird would forage $\sim$ 15 times per h, in agreement with the actual value of 14–18. This result is reasonable in view of the mean meal size of 70 μ l measured in the field (F.L.C. and M. A. Hixon, unpublished observations). Crop emptying is probably limited by the times required for acidification of the stomach contents and/or intestinal nutrient absorption. Thus, despite external appearances, hummingbirds may be energy maximizers, taking in energy as rapidly as their digestive processes permit.

We thank David Solberg, Meredith Hughes, Dr David Mertens and Edith Read for assistance, and Don Powers for providing some birds. This work was supported by NIH grant GM14772 and NSF grant DEB8022575.

Received 1 November; accepted 3 December 1985.

1. Hainsworth, F. R. & Wolf, L. L. *Comp. Biochem. Physiol.* **42A**, 359–366 (1972).
2. Hainsworth, F. R. *J. comp. Physiol.* **88**, 425–431 (1974).
3. Stiles, F. G. *Science* **173**, 818–821 (1971).
4. Hixon, M. A., Carpenter, F. L. & Paton, D. C. *Am. Nat.* **122**, 366–391 (1983).
5. Karasov, W. H., Phan, D., Diamond, J. M. & Carpenter, F. L. *Auk* (in the press).
6. Castle, E. J. *Br. J. Nutr.* **10**, 15–23 (1956).
7. Walsburg, G. E. *Condor* **77**, 169–174 (1975).
8. Karasov, W. H., Rosenberg, L., Petrossian, E. & Diamond, J. M. *J. comp. Physiol.* (in the press).
9. Sibley, R. M. in *Physiological Ecology: An Evolutionary Approach to Resource Use* (eds Townsend, C. R. & Calow, P.) 109–142 (Sinauer, Sunderland, 1983).
10. Karasov, W. H. & Diamond, J. M. *J. comp. Physiol.* **152**, 105–116 (1983).
11. Karasov, W. H., Buddington, R. K. & Diamond, J. M. in *Proc. 1st int. Congr. comp. Physiol. Biochem.* Vol 2 (Springer, Berlin, in the press).
12. Karasov, W. H., Solberg, D. H. & Diamond, J. M. *Am. J. Physiol.* **249**, G271–283 (1985).
13. Hixon, M. A. *Am. Nat.* **119**, 596–599 (1982).
14. Wolf, L. L. & Hainsworth, F. R. *Ecology* **52**, 980–988 (1971).
15. Ewald, P. W. & Carpenter, F. L. *Oecologia* **31**, 277–292 (1978).

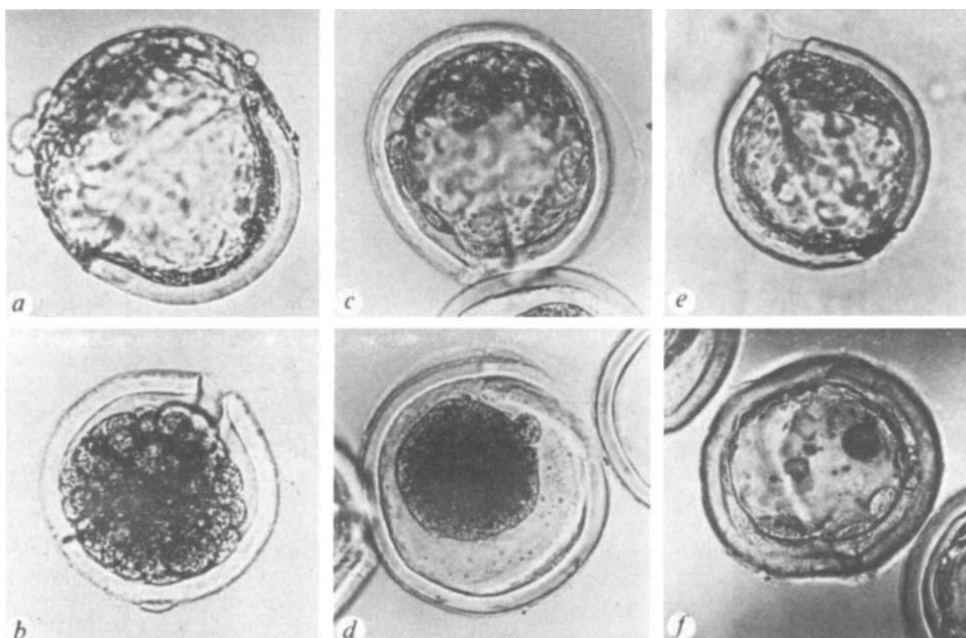
Nuclear transplantation in sheep embryos

S. M. Willadsen

AFRC Institute of Animal Physiology, Animal Research Station, 307 Huntingdon Road, Cambridge CB3 0JQ, UK

Nuclear transplantation and cell fusion techniques have proved valuable for embryological studies in several non-mammalian animal species¹. More recently these procedures have been used successfully in small laboratory mammals, notably the mouse, to investigate the ability of nuclei and cytoplasm from various sources to produce viable embryos when combined^{2–6}. The use of a similar approach to study the developmental biology of large domestic animals presents a number of technical and practical difficulties, and so far there has been no report of attempts to perform nuclear

Fig. 1 Micromanipulated embryos after culture in sheep oviducts, but still embedded in agar. $\sim \times 250$. *a, b*, 'Nuclear-transfer' embryos, cultured for 5 days. These embryos were produced by fusing two blastomeres from an 8-cell embryo with the 'enucleated' (*a*) and the 'nucleated' half (*b*) of an unfertilized egg: *a* was a fully expanded blastocyst containing ~ 120 (nucleated) cells, while *b* contained only a few nucleated cells centrally, the rest of the 'cells' being fragments without nuclei. *c, d*, 'Half' embryos cultured for 5 days. These embryos were produced by dividing a pronuclear embryo (~ 10 h after sperm penetration) in such a way that one half contained both pronuclei (*c*). About two-thirds of such pronuclear 'half' embryos developed into normally organized blastocysts; *c* was an expanding blastocyst containing ~ 60 cells, while *d*, the half without pronuclei, had remained a single cell. *e, f*, 'Single blastomere' embryos. *e*, An expanding blastocyst which developed during 4½ days of culture from a blastomere isolated from a 2-cell embryo. This embryo contained ~ 60 cells. *f*, A vesicle which developed during 3 days of culture from a single blastomere from an 8-cell embryo. This embryo contained ~ 20 cells, but had no proper inner cell mass. Three-quarters or more of 'single blastomere' embryos will continue to develop at a normal rate up to the time of blastulation in the sheep oviduct. Blastocysts of the types shown in *a, c* and *e* appear to have the same viability as ordinary sheep blastocysts, while embryos of the type shown in *f* are rarely viable.



transplantation in sheep embryos. Here I describe such a procedure and its use to investigate the development of embryos in which whole blastomeres from 8- and 16-cell embryos were combined with enucleated or nucleated halves of unfertilized eggs. The procedure involves bisection of single-cell eggs in a medium containing cytochalasin; fusion of egg halves with single blastomeres, induced using Sendai virus or an electrofusion apparatus; and embedding in agar, followed by culture of the reconstituted embryos in the ligated oviducts of ewes in dioestrus. I show that fully viable embryos may be obtained by this procedure.

Eggs and embryos were collected from Welsh Mountain $\times$ Cheviot ewes. A few hours before expected onset of oestrus (day 0), donor ewes received 500 IU human chorionic gonadotropin (hCG; Chorulon, Intervet Ltd) intravenously. Unfertilized eggs were recovered from unmated donors 30–33 h after hCG administration, while 8- and 16-cell embryos were collected on days 3 and 4, respectively, from donor ewes which had been inseminated with Suffolk ram semen. Eggs and embryos were collected, stored, manipulated and transferred in an enriched phosphate-buffered saline medium (PBS)⁷ and kept at room temperature ($\sim 20^\circ\text{C}$) except where stated otherwise. The procedures used for follicular stimulation and oestrus induction in prospective donor ewes, and the surgical techniques used for collection and transfer of oocytes and embryos have been described in detail elsewhere⁸.

Using a fine glass needle, a wide slit was made in the zona

pellucida over the polar body of unfertilized eggs, and the eggs were placed in PBS containing cytochalasin B ($5\text{ }\mu\text{g ml}^{-1}$; Sigma). After about 1 h the polar body and the adjacent ooplasm were sucked into a fine-bore pipette ($\sim 30\text{ }\mu\text{m}$ tip diameter) and removed. In this way 90% or more of the eggs could be divided into two fragments with intact cell membranes, each containing approximately half of the ooplasm. Preliminary studies in which halved oocytes were stained with acridine orange and examined by fluorescence microscopy showed that $\sim 75\%$ of those egg halves that were removed with the polar body contained the oocyte metaphase II chromosomes ('nucleated' egg halves), while the other halves from the same oocytes contained no nuclear structures ('enucleated' egg halves). Single blastomeres isolated from 8- and 16-cell embryos were inserted into zonae pellucidae containing either 'enucleated' or 'nucleated' egg halves. A Leitz micromanipulator and two De Fonbrune micro-syringes were used for these micromanipulations. The glass instruments and techniques used in such procedures have been described in detail previously⁹.

When Sendai virus was to be used to induce cell fusion, the isolated blastomeres were placed in a suspension of inactivated Sendai virus ($\sim 1,000$ haemagglutination units per ml PBS with cytochalasin) for 2 min immediately before being combined with egg fragments as described above. (For discussions of the use of Sendai virus for fusion of embryonic cells see refs 2, 10, 11.) When electrofusion was to be used, the zonae pellucidae con-

Table 1 Development of reconstituted embryos in sheep oviducts

Egg half Blastomere donor	'Enucleated'			'Nucleated'	
	8-cell	8-cell	16-cell	8-cell	16-cell
Fusion method*	S	E	E	E	E
Embryos cultured (%)	24 (100)	76 (100)	29 (100)	35 (100)	19 (100)
Blastulating	8 (33.3)	32 (42.1)	14 (48.3)	4 (11.4)	1 (5.3)
Several cleavages but no blastulation	8 (33.3)	17 (22.4)	6 (20.7)	10 (28.6)	12 (63.1)
One cleavage only	—	2 (2.6)	1 (3.4)	6 (17.1)	2 (10.5)
No cleavage	—	21 (27.6)	4 (13.8)	15 (42.9)	3 (15.8)
Lysis	8 (33.3)	4 (5.3)	4 (13.8)	—	1 (5.3)

* S, Sendai virus-induced fusion; E, electrofusion.

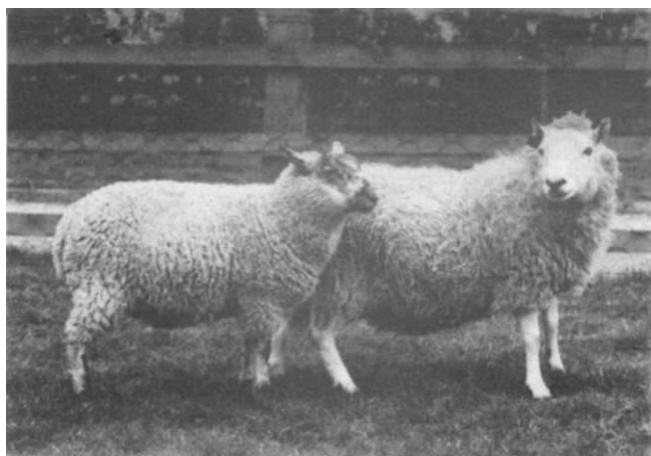


Fig. 2 Lamb (Suffolk crossbred) produced by combining a single blastomere from an 8-cell embryo (Suffolk crossbred) with the 'enucleated' half of an unfertilized Welsh $\times$ Cheviot egg. The 8-cell embryo used as 'nucleus donor' was deep-frozen on 11 November 1979 and thawed on 28 March 1984, the day of the nuclear-transfer experiment.

taining blastomere and egg half were placed in a solution of 0.3 M mannitol, 0.1 mM MgSO_4 and 0.05 mM CaCl_2 in water for 15–30 min. They were then transferred, in the same medium, to the chamber (electrode distance 200 μm) of an electrofusion apparatus (Zimmerman; GCA, Chicago) and exposed to the following fusion conditions: cell alignment (600 kHz, 6 V for 5–10 s), followed by three fusion pulses of 15 V d.c., each lasting 0.1 ms, at 0.1-s intervals; after the fusion pulses, the alignment voltage was reduced from 6 V to 0 V over 1 min (for a review of the use of electrofusion see ref. 12). All embryos were then incubated at 37 °C in PBS containing cytochalasin B. In the Sendai virus-induced group, about 50% of the paired cells fused and this took up to 4 h. By contrast, about 90% of the blastomeres had fused with the egg fragments within 1 h in the electrofusion group.

Those embryos in which cell fusion had occurred were placed in PBS at room temperature. They were then embedded in agar and transferred to ligated oviducts of ewes in dioestrus as described previously¹³. After $4\frac{1}{2}$ to $5\frac{1}{2}$ days, they were recovered from the ewes and examined as fresh specimens. Some of those which had developed into normally organized blastocysts were transferred to recipient ewes on day 6 or 7 of their oestrous cycle. The rest were examined by phase-contrast microscopy after fixation in acetic ethanol (1:3) and staining with 1% Lacmoid in 45% acetic acid.

The development of reconstituted embryos during culture in sheep oviducts is summarized in Table 1. A substantial proportion of the embryos in the 'enucleated' groups developed into early or expanding blastocysts whether cell fusion had been induced electrically or with Sendai virus. Both with respect to the time of blastulation and the cell number at the early blastocyst stage, these embryos behaved like pronuclear eggs rather than single blastomeres from 8- or 16-cell embryos (compare with Fig. 1). Also, the potential for further development was much greater in blastocysts produced in the present experiments than that of vesicular forms developing from single blastomeres from 8-cell sheep embryos, of which less than 10% (2/31) gave rise to lambs¹⁴. In the present experiments, three of the four blastocysts transferred to recipient ewes during the 1983–84 breeding season developed into full-term lambs. All three were produced from reconstituted embryos in which a blastomere from an 8-cell embryo was combined with an 'enucleated' half of an unfertilized egg, and, as expected, all three lambs were of the 8-cell variety (Suffolk crossbred, Fig. 2). Two of the lambs were 'identical' twins. During the current breeding season a

number of pregnancies have been established after transfer of blastocysts produced by combining single blastomeres from 16-cell embryos with 'enucleated' egg halves. Three ewes, each of which had received two blastocysts from a group of six derived from the same 16-cell embryo, were killed on day 60. Each of the three recipients was found to carry one apparently normal fetus.

The present experiments have demonstrated that in the sheep, fully viable embryos can be produced by combining a (nucleus of a) blastomere from an 8-cell embryo with approximately half of the cytoplasm of an unfertilized egg. Furthermore, the results to date strongly indicate that even at the 16-cell stage, the blastomere nucleus remains totipotent. Although these conclusions are of a qualitative rather than a quantitative nature, it seems reasonable to suggest that a firm basis has been established for further experiments involving nuclear transplantation in large domestic species. The most obvious practical aim of such experiments will be to define conditions which will allow large-scale cloning of domestic animals. This would appear to be a realistic aim in view of the present results and those of other studies which have shown that it is possible to obtain blastocysts when blastomeres from cleaving reconstituted embryos are themselves used as nucleus donors (unpublished observation).

Received 18 March; accepted 4 October 1985.

1. McLaren, A. *Nature* **309**, 671–672 (1984).
2. McGrath, J. & Solter, D. *Science* **220**, 1300–1302 (1983).
3. Surani, M. A. H. & Barton, S. C. *Science* **222**, 1034–1036 (1983).
4. Surani, M. A. H., Barton, S. C. & Norris, M. L. *Nature* **308**, 548–550 (1984).
5. Barton, S. C., Surani, M. A. H. & Norris, M. L. *Nature* **311**, 374–376 (1984).
6. McGrath, J. & Solter, D. *Science* **226**, 1317 (1984).
7. Whittingham, D. G. *Nature* **233**, 125–126 (1971).
8. Willadsen, S. M. in *The Management and Diseases of Sheep* (ed. British Council) 69–85 (Commonwealth Agricultural Bureaux, Slough, 1980).
9. Willadsen, S. M. in *Mammalian Egg Transfer* (ed. Adams, C. E.) 185–210 (CRC Press, Boca Raton, 1982).
10. Graham, C. F. in *Proc. 3rd Karolinska Symp. on Research Methods in Reproductive Endocrinology* Stockholm, 154–165 (1971).
11. McGrath, J. & Solter, D. *J. exp. Zool.* **228**, 355 (1983).
12. Zimmerman, V. & Vienen, J. J. *Membrane Biol.* **67**, 165–182 (1982).
13. Willadsen, S. M. *Nature* **277**, 298–300 (1979).
14. Willadsen, S. M. *J. Embryol. exp. Morph.* **65**, 165–172 (1981).

Structural alteration of hair cells in the contralateral ear resulting from extracochlear electrical stimulation

H. C. Dodson*, J. R. Walliker†‡, S. Frampton†, E. E. Douek§, A. J. Fourcin‡ & L. H. Bannister*

* Department of Anatomy, Guy's Campus, United Medical and Dental Schools of Guy's and St Thomas' Hospitals, and

† Department of Clinical Physics and Bioengineering and § Ear, Nose and Throat Department, Guy's Hospital, London SE1 9RT, UK

‡ Department of Phonetics and Linguistics, University College London, 4 Stephenson Way, London NW1 7EH, UK

Chronic electrical stimulation of the auditory nerve in patients with profound sensori-neural deafness is becoming increasingly routine. Therefore, it is important to understand more about the long-term consequences of this procedure. Hitherto, structural studies in animals after electrocochlear stimulation have concentrated on the stimulated cochlea^{1–6}. Here we have examined the effects of unilateral extracochlear electrical stimulation on the spiral organ of both the ipsilateral and contralateral ears of the mature guinea pig, and have found alterations in the structure of the outer hair cells and their efferent nerve terminals in the contralateral as well as the ipsilateral cochlea. This is the first evidence for a structural influence of efferent activity on the cochlea. Although the importance of the efferent system, consisting of the crossed and uncrossed olivo-cochlear bundles, is well estab-

lished in providing central control of the sensory pathways⁷, its exact role in hearing is incompletely understood. However, it is known that the outer hair cells and their efferent innervation are important in their contribution to inner hair cell responses^{8,9} and in modulating the micromechanics of the whole cochlea^{10,11}. These efferent functions now appear to be related to an important part of cochlear morphology, and are also relevant to our understanding of cochlear neurobiology, normal development and the management of hearing disability in both adult and child.

Using the method of Aran and Erre^{12,13}, 15 guinea pigs (albino Dunkin Hartley and a pigmented strain derived from a colony maintained by the Royal College of Surgeons, London), each weighing 400–500 g, were implanted with an electrode ball positioned at the round window in the middle ear. The electrode material was either platinum or platinum containing 10% iridium. The electrode wire was passed subcutaneously to a head connector. Stainless-steel screws in the skull served as the reference and indifferent electrodes. The electrical stimulus was delivered via a thin counter-weighted coaxial cable, allowing the animal unrestricted movement. The animals were stimulated for 12 h per day for periods of up to 20 weeks (1,100 h of stimulation) with a charge-balanced square-wave alternating current at 100 Hz. Current levels ranged from 25 to 91 μA r.m.s., depending on the individual tolerance of each guinea pig. The direct-current component of the stimulus was negligible at <30 pA. The compound action potential (CAP) of the auditory evoked response (AER) and thresholds to click stimuli delivered at a rate of 10 s^{-1} with a bandwidth of 100 Hz to 6 kHz, ± 10 dB, were recorded from the stimulating electrode in un-anaesthetized guinea pigs before stimulation (7 days after implantation) and at regular intervals during the electrical stimulation period. The 12-h stimulation period of the 24-h cycle was interrupted to make AER measurements. At the end of the experiment the animals were killed either immediately after electrical stimulation or at periods of up to 5 days after disconnection, and the spiral organ from each cochlea was examined by transmission electron microscopy. As controls, 14 (12 albino, 2 pigmented) age-matched implanted but unstimulated animals, and 10 (8 albino, 2 pigmented) unimplanted animals were also maintained in the same environment and were examined in the same way.

Ipsilateral AER audiometry in stimulated guinea pigs showed no appreciable change in threshold or latency during the experiment, with thresholds remaining constant at a sound pressure level of 25–30 dB (r.m.s. impulse time weighting¹⁴), as measured in a quiet room. These were similar to those recorded periodically from the implanted, unstimulated controls. AER measurements made at different times of day showed no significant variations in threshold or latency. Six animals (three pigmented, three albino) stimulated for periods of <200 h at current levels of 25 μA r.m.s. demonstrated few cytological changes in either the stimulated or contralateral cochlea (Fig. 1a). However, in eight guinea pigs (five albino, three pigmented) stimulated for prolonged periods (maximum 1,100 h), we found ultrastructural alterations in large numbers of outer hair cells in the basal turns of both cochleae. These alterations were more marked in the contralateral than the ipsilateral cochlea, although of the same general pattern. In three guinea pigs (one pigmented, two albino) all the outer hair cells in the basal turn were severely altered and in the remaining animals (two pigmented, three albino) approximately 50% of the outer hair cells in the basal turn were affected, although less severely. Similar changes were seen in animals killed both immediately after the end of stimulation and up to 5 days after disconnection. We observed fusion and bending of the stereocilia of the outer hair cells, folding of the plasma membrane and swelling of the smooth endoplasmic reticulum (Fig. 1b). The efferent endings of the olivo-cochlear tract situated at the bases of the outer hair cells stained densely, contained increased numbers of synaptic vesicles, and were irregular in profile compared with controls (Fig. 1c, d). The mitochondria within these endings were often grossly vacuolated

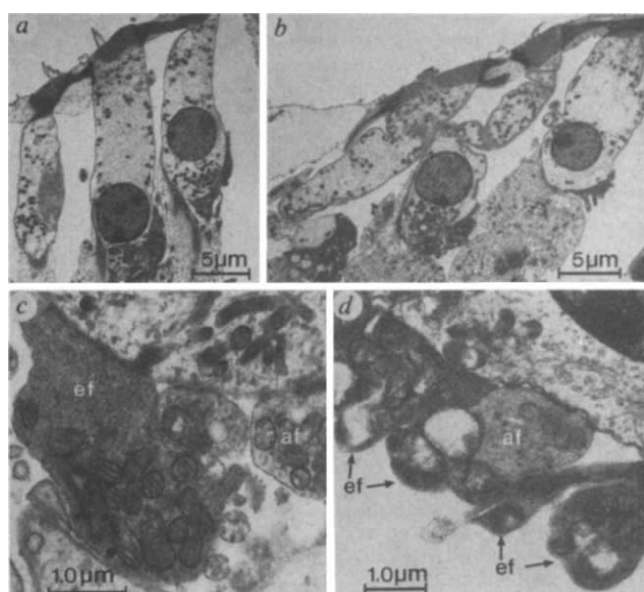


Fig. 1 Electron micrographs of outer hair cells and synaptic terminals from guinea pigs. *a*, Sensory cells from the basal turn of a guinea pig cochlea (contralateral to the electrode), stimulated at 25 μA r.m.s. for 93 h. The outer hair cells show no obvious abnormalities and their basal synapses also appear normal. *b*, Cells from the same region on the contralateral side of a guinea pig stimulated at 91 μA r.m.s. for 1,100 h with a (unilateral) extracochlear electrode; the outer hair cells show various signs of collapse and their stereocilia are bent. *c*, A group of synaptic endings at the base of an unstimulated outer hair cell, including a large efferent (ef) and an afferent (af) terminal. *d*, A similar section from the contralateral ear of an electrically stimulated animal showing irregular, densely stained, vacuolated efferent endings and a normal afferent terminal; abbreviations as in *c*.

although those in surrounding cells were normal. The cells of Deiters also demonstrated irregularities of shape and an increase in smooth endoplasmic reticulum. Inner hair cells and afferent endings appeared to be unchanged. Unstimulated animals, both implanted and unimplanted, did not show appreciable alterations in structure.

The results presented here show that although chronic electrical stimulation under the conditions described did not alter the structure of the afferent cochlear nerve fibres and ganglion cells, it did result in changes to the outer hair cells and efferent terminals in the basal turns, particularly of the contralateral cochlea. So far, we have been unable to find physiological correlates of the morphological damage; experiments are in progress to explore this aspect further.

It is known that electrical stimulation of the crossed olivo-cochlear pathway in the medulla causes contralateral inhibitory effects on hair cell responses to acoustic stimuli^{15,16}. Similar (but ipsilateral) effects occur with unilateral extracochlear electrical stimulation¹⁷. It is not yet known whether contralateral effects can be induced by this technique although unilateral acoustic stimulation has been shown to affect auditory responses on the contralateral side¹⁸.

The results reported here demonstrate that chronic activation of the crossed and uncrossed olivo-cochlear pathways leads to changes within the outer hair cells. It is possible that this effect is reversible although the alterations were still present 5 days after the end of stimulation. The fact that these changes are restricted to the basal turns of the cochlea indicates that they are induced by proximity (in the ipsilateral ear) to the electrode rather than by activation of the whole cochlear nerve. The similar distribution of damage in the contralateral cochlea suggests that the spatial organization of the efferent system on the two sides is related.

Our finding that electrical stimulation of the afferent pathways

of the cochlear nerve in one ear can result in alterations to the organ of Corti in the contralateral ear, associated with the sustained activity of the crossed olivo-cochlear pathway, is important for auditory neurobiology. As there now exist considerable possibilities for the surgical provision of electrocochlear hearing aids for profoundly and totally deaf adults and children by intracochlear and extracochlear methods, it is essential to have a more comprehensive understanding of the long-term implications of these procedures.

This work was supported by the MRC (UK). We thank Dr Cazals and J.-P. Erre for their tuition in the technique of round window implantation. The programmable multi-channel stimulator was constructed by H. Carson of the Department of Clinical Physics and Bioengineering, Guy's Hospital. We also thank M. Blagg of the same department for typing the manuscript.

Received 18 November 1985; accepted 3 January 1986.

1. Sutton, D., Miller, J. M. & Pfingst, B. E. *Ann. Otol. Rhinol. Lar.* **89**, Suppl. 66, 11-14 (1980).
2. Sutton, D. & Miller, J. M. *Ann. Otol. Rhinol. Lar.* **92**, 53-58 (1983).
3. Duckert, L. G. & Miller, S. M. *Ann. Otol.* **91**, 33-40 (1982).
4. Duckert, L. G. & Miller, J. M. *Acta oto-lar. Suppl.* **411**, 28-37 (1984).
5. Schindler, R. A. & Merzenich, M. M. *Ann. Otol. Rhinol. Lar.* **83**, 202-216 (1974).
6. Leake Jones, P. A., Walsh, S. M. & Merzenich, M. M. *Ann. Otol. Rhinol. Lar.* **90**, Suppl. 82, 6-8 (1981).
7. Desmedt, J. E. in *Handbook of Sensory Physiology* Vol. 5, Pt 2 (eds Keidel, W. D. & Neff, W. D.) 219-246 (Springer, Berlin, 1975).
8. Harrison, R. V. & Evans, E. F. in *Inner Ear Biology* (eds Portmann, M. & Aran, J. M.) 105-124 (Inserm, Paris, 1977).
9. Carlier, E. & Pujol, R. *Neurosci. Lett.* **28**, 101-106 (1982).
10. Mountain, D. C. *Science* **210**, 71-72 (1980).
11. Siegel, J. H. & Kim, D. O. *Hearing Res.* **6**, 171-182 (1982).
12. Aran, J. M. & Erre, J.-P. in *Audiological Investigation: the Scientific and Technological Basis* (ed. Beagley, H. A.) 233-261 (Clarendon, Oxford, 1979).
13. Aran, J. M. & Cazals, Y. in *Evoked Electrical Activity in the Auditory Nervous System* (eds Naunton, R. F. & Fernandez, C. P.) 239-257 (Academic, New York, 1978).
14. *Int. Electrotech. Commn. Publ.* 651.
15. Konishi, T. & Slepian, J. Z. *J. acoustic. Soc. Am.* **49**, 1762-1769 (1971).
16. Brown, M. C. & Nuttall, A. L. *J. Physiol., Lond.* **354**, 625-646 (1984).
17. Rajan, R. & Johnstone, B. M. *Hearing Res.* **12**, 405-417 (1983).
18. Feldman, H. *Audiology* **10**, 138-144 (1971).

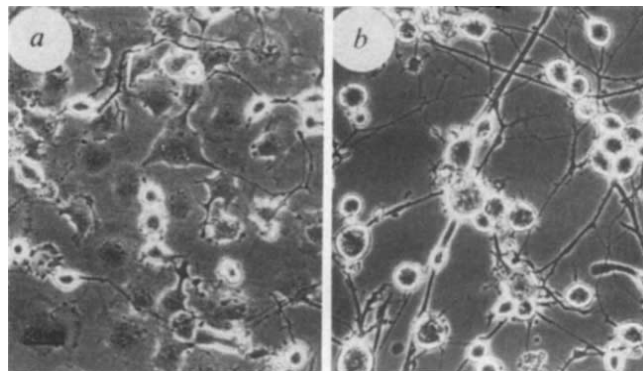


Fig. 1 Cell morphology during the ecto-protein kinase assays. *a*, NG108-15 cells were seeded in individual wells of 96-well plates and grown in a chemically defined, serum-free medium until reaching 80-90% confluency (6 days). The medium was prepared essentially as described previously²⁰, except that the concentrations of added supplements were: 25 $\mu\text{g ml}^{-1}$ insulin, 50 $\mu\text{g ml}^{-1}$ transferrin, 2.5 $\mu\text{g ml}^{-1}$ linoleic acid and 25 mM HEPES (pH 7.4). *b*, Cells were grown as in *a* and then treated with 1 mM dibutyryl cAMP for 4 days. Medium and drug were changed daily. The cells were photographed in the wells in which they were subsequently assayed for protein phosphorylation activity. Note that after dibutyryl cAMP treatment (*b*), all the cells appear with the morphological characteristics of differentiated NG108-15 cells^{10,11}, including long, thick neurites, varicosities and oval, shining cell bodies. Such cells can also be seen in *a*; in cultures maintained in the chemically defined medium for 8-12 days, cells with this morphology constituted 30-50% of the total population. For the ecto-protein kinase assays shown in Figs 2-4, we used cells maintained in the chemically defined medium for 8-10 days. Microscopic examination after the reactions has verified that their morphology did not change during the assay. The clonal cell line used in these studies was provided by Dr M. Nirenberg.

Ecto-protein kinase activity on the external surface of neural cells

Yigal H. Ehrlich*†, Theodore B. Davis*, Elisabeth Bock†, Elisabeth Kornecki & Robert H. Lenox

Neuroscience Research Unit, Department of Psychiatry and

* Department of Biochemistry, University of Vermont, Burlington, Vermont 05405, USA

† The Protein Laboratory, University of Copenhagen, DK-2200 Copenhagen-N, Denmark

ATP is secreted in association with neurotransmitters at certain synapses and neuromuscular junctions^{1,2}. Extracellular ATP is known to exert potent effects on the activity of cells in the nervous system, where it can act as a neurotransmitter or as a modulator regulating the activity of other neurohormones³⁻⁵. We have suggested that such modulation may involve the activity of extracellular protein phosphorylation systems⁶. It is well known that intracellular protein kinases are important in the regulation of various neuronal functions⁶⁻⁸, but protein kinases which use extracellular ATP to phosphorylate proteins localized at the external surface of the plasma membrane (ecto-protein kinases) have not been demonstrated in neuronal cells. Here we present direct evidence for the existence of an ecto-protein kinase and demonstrate endogenous substrates for its activity at the surface of intact neural cells. The phosphorylation of one of these surface proteins is selectively stimulated during cell depolarization. In addition, neuronal cell adhesion molecules (N-CAMs) appear to be among the substrates of ecto-protein kinase activity. These results suggest a role for surface protein phosphorylation in regulating specific functions of developing and mature neurones.

Conclusive evidence for the existence of an ecto-protein kinase and identification of the endogenous substrates for its activity require that the studies be carried out with a homogeneous population of intact, viable cells⁹. To meet these requirements, we investigated here a hybrid cell line of neuronal origin, clone NG108-15 (refs 10, 11). To exclude the possibility that phosphorylating activity measured with intact cultured cells arises from a contaminating serum protein, or resides in membrane fragments formed when some cells break during the collection procedure, we have studied NG108-15 cells which were seeded into individual wells of 96-well plates and grown in a chemically defined, serum-free medium for 8-10 days (see Fig. 1 legend). During this period the medium was changed daily and the cells reached 80-90% confluency (see Fig. 1*a*). After rinsing the cells twice with a reaction buffer mimicking a physiological extracellular environment (see Fig. 2 legend) they were assayed for protein phosphorylation by extracellular [γ -³²P]ATP while still attached to the bottom of each well. The reaction conditions did not affect cell morphology or viability as determined by microscope examination, trypan blue exclusion or continued growth of the cells for the next 2-4 days.

The products of endogenous protein phosphorylation activity in intact NG108-15 cells incubated with extracellular [γ -³²P]ATP are shown in Fig. 2. The major protein substrates for this activity migrated in SDS gels with apparent relative molecular masses (M_r) of >200,000, 190,000, 134,000, 120,000, 113,000, 94,000, 78,000 and 47,000 (Fig. 2, lane *a*). Extraction with chloroform/methanol, cold and hot sodium hydroxide, hot trichloroacetic acid and hydroxylamine provided evidence that the incorporated ³²P was bound to proteins in phosphomonoester linkages, as expected from products of protein kinase activity^{6,12,13}. Treatment of slab gels with 1 M KOH at 55 °C for 1 h indicated that the phosphorylated proteins do not contain ³²P-tyrosine (data not shown). The surface location of the protein kinase phosphorylating these proteins was demon-

† To whom reprint requests should be addressed at the Department of Psychiatry.

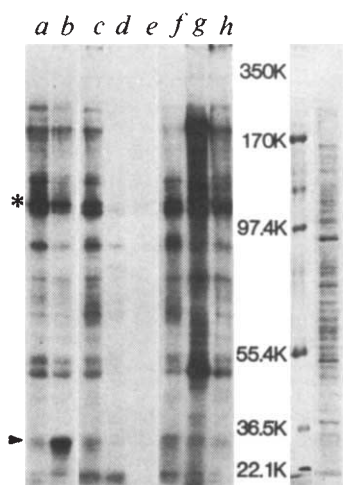


Fig. 2 Endogenous protein substrates of ectokinase activity in intact NG108-15 cells. *a*, Reaction in incubation buffer only; *b*, reaction with attached cells plus 1.0 μ g α -casein added to the incubation medium (see arrow); *c*, reaction initiated after 10 min preincubation of the cells (37 °C) in the reaction medium; *d*, reaction initiated after 10 min preincubation of attached cells with 0.01% trypsin at 37 °C; *e*, re-incubation for 10 min at 37 °C with 0.01% trypsin, added after the reaction; *f*, reaction in buffer supplemented with 4 mM $MgCl_2$; *g*, reaction in buffer supplemented with 4 mM $MnCl_2$; *h*, 10 μ M APPNHP was added 10 s before the ATP. Shown on the right are the M_r markers²² and the pattern of protein staining by Coomassie blue, which appeared the same for all the lanes. Results were obtained with 6–12 separate batches of cells assayed in passages 22–30. Each reaction condition was tested in duplicate or triplicate. Asterisk denotes the protein band referred to in the text as the major substrate of ecto-protein kinase activity. Arrow marks the position of added casein in lane *b*.

Methods. Cells were grown in 96-well plates as described in Fig. 1*a*, maintained in the chemically defined medium (with medium changed daily) for 10 days, and then rinsed twice in a reaction buffer containing 145 mM NaCl, 5 mM KCl, 0.8 mM $MgCl_2$, 1.8 mM $CaCl_2$, 20 mM glucose and 25 mM HEPES buffer, adjusted to pH 7.4 with NaOH. All the reactions shown here (final volume 100 μ l) were initiated by adding 1 μ M [γ -³²P]ATP (10 μ Ci per well) to attached cells, and terminated by adding SDS stop solution²¹ after 10 min incubation with gentle shaking at 37 °C. Samples were resolved on 7–14% exponential polyacrylamide gel gradients and autoradiographed for 6 days as described elsewhere^{21,22}.

strated by the ability of attached, rinsed, intact cells to phosphorylate the exogenous protein substrate, casein, added to the extracellular medium (Fig. 2, lane *b*). Casein competes with endogenous protein substrates for the ectokinase, as suggested by the decreased phosphorylation of cell proteins observed in the presence of casein (compare Fig. 2 lanes *a* and *b*). The localization of the endogenous protein substrates at the cell surface was tested by treating intact NG108-15 cells with very low concentrations of trypsin (Fig. 2; compare lanes *d*, *e* with lane *c*). Preincubation of intact, attached cells with 0.01% trypsin eliminated the phosphorylation activity (lane *d*), whereas treatment with 0.01% trypsin after the reaction removed the ³²P-phosphate incorporated into cellular proteins from extracellular [γ -³²P]ATP (lane *e*). The conclusion that this endogenous phosphorylation system is located at the cell surface was confirmed, as preincubation of intact cells with a specific antibody directed against surface proteins (see Fig. 4*a* legend) selectively inhibits the phosphorylation of these proteins by extracellular ATP.

Addition of cyclic AMP, cGMP, calmodulin, exogenous phospholipids or tetradecanoyl phorbol 13-acetate to ectokinase assays in intact NG108-15 cells has no significant effects on the phosphorylation patterns shown in Fig. 2*a* or 2*f*. Inclusion of 4 mM $MnCl_2$ has selective stimulatory effects on the phosphory-

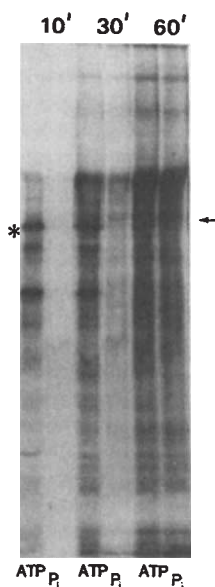


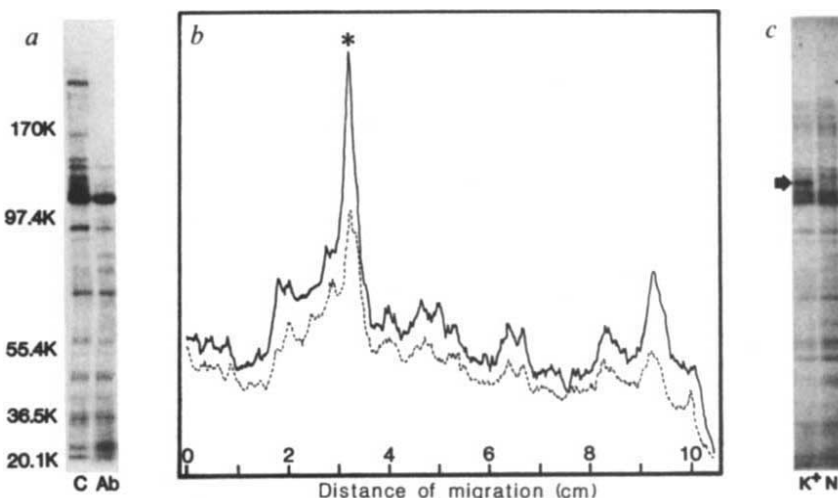
Fig. 3 Time course of protein phosphorylation in intact NG108-15 cells by radiolabelled ATP and inorganic phosphate. Cells were grown and maintained in the chemically defined medium for 8 days and then assayed as described in Figs 1*a* and 2 legends. In the lanes marked ATP, reactions were initiated by adding 2 μ M [γ -³²P]ATP (20 μ Ci per well) and proceeded for the times indicated above the autoradiogram. In the lanes marked P_i , an equivalent amount of ³² P_i (20 μ Ci per well) was added to adjacent wells of the same plate (2 μ M non-radiolabelled ATP was added with the ³² P_i to ensure identical incubation conditions). Reaction products were solubilized in SDS and resolved on a 7% polyacrylamide gel. Asterisk indicates the major substrate of ecto-protein kinase, which is not labelled by incubating cells with ³² P_i (compare 60 min with 10 min labelling). Arrow on the right points to a major phosphorylated band in cells labelled with ³² P_i , which migrates more slowly than the one marked by an asterisk.

lation of many (though not all) of the endogenous proteins by extracellular ATP (Fig. 2, compare lanes *g* and *f*). A slowly degradable analogue of ATP has selective inhibitory effects (lane *h*), whereas dilution of the added [γ -³²P]ATP with increasing concentrations of non-radiolabelled ATP results in gradual decrease of detectable bands in the autoradiograph (data not shown). In parallel assays we found that none of these agents has any effect on the phosphorylation pattern obtained with cells in which intracellular ATP was labelled for 10–60 min (see Fig. 3) or for 4 h with labelled inorganic phosphate. This provides further evidence for the extracellular location of the activity shown in Fig. 2.

Incorporation of ³²P-phosphate from extracellular [γ -³²P]ATP into the major protein substrate of ectokinase activity (with an apparent M_r of 120,000, marked by an asterisk in Fig. 2) was seen within the shortest reaction time measured (30 s), and then increased up to at least 10–20 min (note asterisk in Fig. 3). In contrast, when cells were incubated with equivalent amounts of ³² P_i to label intracellular ATP pools, protein phosphorylation was observed only after 30–60 min incubation (Fig. 3), and then produced a pattern different from that obtained with [γ -³²P]ATP (compare, for example, the phosphorylated band indicated by an arrow with the one indicated by an asterisk in Fig. 3). Thus, the phosphorylation of proteins seen in intact NG108-15 cells incubated with extracellular [γ -³²P]ATP for up to 20 min cannot be attributed to ³² P_i liberated by ecto-ATPases, which may then label intracellular ATP pools and, in turn, intracellular proteins. We believe that, taken together, our results provide direct evidence that NG108-15 cells have an ecto-protein kinase activity which phosphorylates specific protein substrates localized at the cell surface.

The extracellular phosphorylation of surface proteins by ATP secreted from neurones could play an important part in cell-cell

Fig. 4 Neuronal functions of surface phosphoproteins. **a**, Phosphorylation of N-CAMs by extracellular ATP. Intact NG108-15 cells were assayed for endogenous protein phosphorylation by extracellular [γ - 32 P]ATP as described in Fig. 2 legend, except that before initiation of the reactions the cells were preincubated for 10 min at 37 °C with control buffer (lane C), or with 10 μ l per well of a 1:10 dilution of purified anti D2-CAM antibody¹⁵ (lane Ab). The results shown are representative of four separate assays, each carried out in triplicate. Similar preincubations with control antibody (anti-albumin); had no significant effects on the phosphorylation pattern. **b**, Phosphorylation of proteins by extracellular ATP in differentiated NG108-15 cells. Cells were grown in a chemically defined medium, differentiated by treatment with dibutyryl cAMP as described for Fig. 1b, and incubated with [γ - 32 P]ATP under the reaction conditions specified for Fig. 2, lane **a**. The autoradiograph was scanned (solid line) by a laser microdensitometer as described previously²². The broken line is a scan of an autoradiograph obtained after electrophoresis of the same amount of protein (10 μ g) from reaction of untreated cells (see Fig. 1a), assayed under the same conditions as the differentiated cells, and resolved on different lanes of the same slab gel. Similar results were obtained in six separate assays, each carried out in quadruplicate. **c**, Ecto-protein kinase activity in depolarized cells. NG108-15 cells were grown in a chemically defined medium, differentiated with dibutyryl cAMP and then incubated with [γ - 32 P]ATP (see Fig. 2 legend) either in the standard reaction medium (lane N), or in a medium in which 50 mM KCl were included in place of 50 mM NaCl (lane K⁺). An autoradiograph obtained after a short exposure time is shown here to demonstrate clearly the effects of KCl (band marked with arrow), and therefore some bands seen in Figs 2 and 4a are very weak. Identical results were obtained in four separate experiments, each carried out in duplicate. Asterisk marks the band also indicated by an asterisk in Figs 2, 3.



interactions. We have initiated this investigation by testing whether neuronal cell adhesion molecules (N-CAMs) (refs 14, 15) serve as substrates for ecto-protein kinase activity in NG108-15 cells. As can be seen in Fig. 4a, addition of a purified anti-D2-CAM antibody¹⁵ to the reaction medium of intact NG108-15 cells incubated with extracellular [γ - 32 P]ATP selectively inhibited the phosphorylation of high- M_r protein substrates of ectokinase activity in these cells. These included proteins with apparent M_r s of 140,000 and 190,000 to >200,000, that have previously been reported for N-CAMs (refs 14, 15). Other affected phosphoproteins may be immunochemically related to N-CAM or their phosphorylation may be sterically hindered by antibody molecules bound to N-CAMs. The ability of the antibody to inhibit this phosphorylation activity without penetrating the cell provides a powerful tool for the investigation of the role of the phosphorylation state of CAMs in neuronal functions studied in culture or *in vivo*.

When NG108-15 cells are treated for 4–10 days with 1 mM dibutyryl cAMP their growth is arrested, they become non-tumorigenic, they differentiate morphologically, physiologically and biochemically and exhibit an increase in many properties characteristic of mature neurones^{10,11}. As reported previously for other neuronal cell lines¹⁶, partial differentiation of NG108-15 cells was observed in cells maintained for 8–10 days in a medium without serum¹⁷ (see also Fig. 1 legend). Such cells produced the pattern of ectokinase activity shown in Fig. 2 and by the dashed line in Fig. 4b. After treatment with dibutyryl cAMP in a chemically defined, serum-free medium, all the NG108-15 cells assumed the differentiated morphology shown in Fig. 1b. As shown in Fig. 4b (solid line), phosphorylation of endogenous protein substrates by ectokinase activity increases significantly after complete differentiation of NG108-15 cells. A twofold increase occurs in the phosphorylation of the major protein substrate (asterisk in Figs 2–4), suggesting that this phosphorylation system is associated with the neuronal functions of differentiated NG108-15 cells.

Stimulation of NG108-15 cells by depolarizing KCl concentrations elicits specific neuronal responses (for example, activation of voltage-sensitive Ca²⁺ channels and evoked acetylcholine release), but the response occurs only in cells differentiated with dibutyryl cAMP and not in untreated cells¹⁰. We have found that inclusion of 50 mM K⁺ (in place of 50 mM

Na⁺) in ectokinase assays with differentiated NG108-15 cells selectively stimulates the phosphorylation of one specific protein component by extracellular ATP (a protein with an apparent M_r of 126,000, see arrow in Fig. 4c). On the other hand, the phosphorylation pattern obtained with untreated cells (see Fig. 2, lane **a**) is not altered significantly by depolarizing KCl. Detailed studies of the unique response shown in Fig. 4c should enable the determination of the role of the phosphorylation of a surface protein by ectokinase activity in specific neuronal functions evoked by depolarization.

We reported previously that extracellular ATP can regulate the uptake of Ca²⁺ ions by intact NG108-15 cells¹⁸; ecto-protein kinase may be involved in this process. It has also been reported that when differentiated NG108-15 cells are co-cultured with myotubes they form functional neuromuscular junctions^{10,19}. NG108-15 cells differentiated in a chemically defined medium therefore represent a *homogeneous* cell population (Fig. 1b) suitable for biochemical studies on the modulatory role of extracellular ATP in Ca²⁺-dependent acetylcholine release⁵. The involvement of ecto-protein kinase and its specific substrates in the regulation of this and other events associated with presynaptic function^{1,2,10,11} can readily be studied in this model system (see Fig. 4). Furthermore, the results reported here provide the basis for investigation of the role of surface protein phosphorylation in mediating the effects of extracellular ATP on various cellular functions in the peripheral and central nervous system.

We thank W. Ohlsson, I. Galbraith and J. Wesson for technical assistance, Dr S. Weiss for discussions, P. Smith and M. LaCasse for typing the manuscript and D. Hardwick for preparing the figures. This work was supported by USAFOSR grant 84-0331, NIH grant MH-35735 and NSF grant BNS82-09265.

Received 19 November; accepted 24 December 1985.

1. Silinsky, E. M. & Hubbard, J. I. *Nature* **243**, 404–405 (1973).
2. Kuroda, Y. *J. Physiol., Paris* **74**, 463–470 (1978).
3. Burnstock, G. *Pharmac. Rev.* **24**, 509–581 (1972).
4. Phillips, J. W. & Wu, P. H. in *Physiology and Pharmacology of Adenosine Derivatives* (eds Daly, J. W., Kuroda, Y., Phillips, J. W., Shimizu, H. & Ui, M.) 219–236 (Raven, New York, 1983).
5. Silinsky, E. M. & Ginsborg, B. L. *Nature* **305**, 327–328 (1983).
6. Ehrlich, Y. H. in *Handbook of Neurochemistry* Vol. 6 (ed. Lajtha, A.) 541–574 (Plenum, New York, 1984).
7. Nestler, E. J. & Greengard, P. *Nature* **305**, 583–588 (1983).
8. Gispen, W. H. & Routtenberg, A. (eds) *Brain-Phosphoproteins* (Elsevier, Amsterdam, 1982).
9. Dicker, P., Heppelard, L. A. & Rozengurt, E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2103–2107 (1980).

10. Nirenberg, M. *et al.* *Science* **222**, 794–799 (1983).
11. Hamprecht, B., Glaser, T., Reiser, G., Bayer, E. & Propst, F. *Meth. Enzym.* **109**, 316–341 (1985).
12. Ehrlich, Y. H. & Routtenberg, A. *FEBS Lett.* **45**, 237–243 (1974).
13. Ehrlich, Y. H., Davis, L. G., Gilfoil, T. & Brunngraber, E. G. *Neurochem. Res.* **2**, 533–548 (1977).
14. Sorkin, B. C., Hoffman, S., Edelman, G. M. & Cunningham, B. A. *Science* **225**, 1474–1478 (1984).
15. Lyles, J. M., Linneman, D. & Bock, E. *J. Cell Biol.* **99**, 2082–2091 (1984).
16. Prasad, K. N. *Regulation of Differentiation in Mammalian Nerve Cells* (Plenum, New York, 1980).
17. Davis, T. B., Kornecki, E., Lenox, R. H. & Ehrlich, Y. H. *Soc. Neurosci. Abstr.* **10**, 196 (1984).
18. Ehrlich, Y. H., Whittemore, S. R., Garfield, M. K., Graber, S. G. & Lenox, R. H. in *Brain-Phosphoproteins* (eds Gispén, W. H. & Routtenberg, A.) 375–396 (Elsevier, Amsterdam, 1982).
19. Nelson, P., Christian, C. & Nirenberg, M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1314–1318 (1978).
20. Botstein, J. in *Functionally Differentiated Cell Lines* (ed. Sato, G.) 155–184 (Liss, New York, 1981).
21. Ehrlich, Y. H., Brunngraber, E. G., Sinha, P. K. & Prasad, K. N. *Nature* **265**, 238–240 (1977).
22. Whittemore, S. R., Graber, S. G., Lenox, R. H., Hendley, E. D. & Ehrlich, Y. H. *J. Neurochem.* **42**, 1685–1696 (1984).

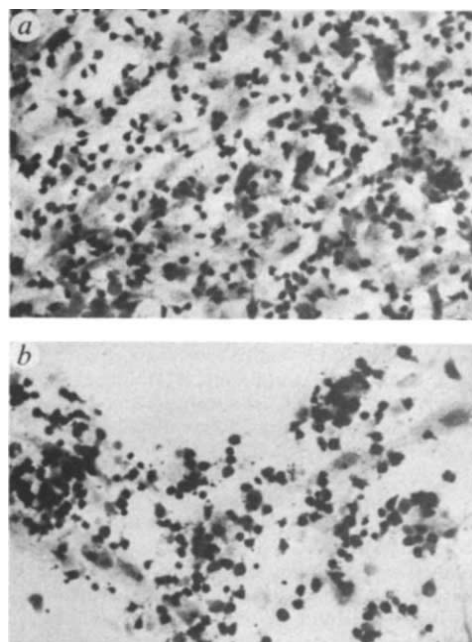


Fig. 1 MBP-specific T lymphocytes (Z1a) cultured on syngeneic astrocyte monolayers in the absence (a) and presence (b) of MBP. Note the intact astrocyte monolayer in the absence of MBP (a). In the presence of MBP, most astrocytes have been lysed, leaving only a fibrillary network. The few remaining cells release blebs and granular material.

Methods. The astrocytes were derived after five passages from cultures of neonatal rat brain cells. Primary cultures were prepared by trypsin dissociation of minced brain matter, which had been trimmed free of adherent meningeal tissue. After five passages, >95% of the monolayer-forming cells were positive for glial fibrillary acidic protein. For interaction with T cells, 1×10^5 pure astrocyte populations were transferred to a 35-mm plastic Petri dish. We added supernatant of Con A-stimulated mouse spleen cells (10%) as a source of Ia-inducing IFN- γ 72 h before the assay. T cells were added (1×10^6 cells per dish) in the presence or absence of $10 \mu\text{g ml}^{-1}$ MBP. After 24 h, the cultures were stained with May-Grünwald-Giemsa.

Ia-restricted encephalitogenic T lymphocytes mediating EAE lyse autoantigen-presenting astrocytes

D. Sun & H. Wekerle

Max-Planck-Gesellschaft, Klinische Forschungsgruppe für Multiple Sklerose, D-8700 Würzburg, FRG

T lymphocytes specific for myelin basic protein (MBP) are responsible for the cellular events leading to autoimmune disease within the central (CNS) and peripheral (PNS) nervous systems. Both in actively induced and T-cell transfer versions of experimental autoimmune encephalomyelitis (EAE) and neuritis (EAN), the autoaggressive T cells are activated outside the nervous system and reach their target tissue via the blood circulation. The target specificity of the autoaggressive T cells is impressive; T-cell lines specific for MBP predominantly home to and affect the white matter of the CNS^{1–3} whereas T cells specific for PNS myelin protein P₂ exclusively infiltrate peripheral nerves^{4,5}. Having penetrated the tight blood tissue barriers, the lymphocytes seem to interact with local cells expressing the relevant autoantigen in an immunogenic form. Although the exact mechanism of target finding and destruction is unknown, studies from our laboratory have shown that astrocytes, a main component of the normal CNS glia, can actively present antigen to specific T cells^{6,7}. This observation suggests that astrocytes are involved in natural immune reactivity within the CNS, and that they may be involved in pathological aberrations, such as in the development of autoimmune lesions. Having studied astrocyte/T-cell interactions in more detail, we have discovered that encephalitogenic T-cell lines recognizing MBP on astrocytes will subsequently proceed to kill the presenting cells. Here we report that astrocyte killing follows the rules governing 'classical' T-cell-mediated cytotoxicity; it is antigen-specific, restricted by antigens of the major histocompatibility complex (MHC) and apparently contact-dependent. Our data suggest that the nature of the recognized antigenic epitope determines whether or not antigen recognition is followed by killing; moreover, killing of antigen-presenting astrocytes seems to be correlated with the capacity to transfer encephalomyelitis to normal syngeneic rats.

When rat T lymphocytes are seeded on syngeneic astrocyte monolayers, they immediately start to migrate to the top of the cell surfaces. They extend uropods, but show no morphological signs of activation such as blast transformation or proliferation. Induction of Ia antigen on the astrocyte membranes by γ -interferon (IFN- γ) does not affect this response pattern. The T cells may persist in such cultures for extended periods, without addition of exogenous growth factors (Fig. 1a). After addition of the relevant antigen (MBP) to Ia⁺ astrocyte cultures, the lymphocyte/astrocyte contacts become much firmer. In contrast to antigen-free cultures, where 50–60% of all seeded lym-

phocytes can be recovered by gently pipetting, practically no viable T cells are released 4 h after addition of MBP. Within 14 h the adherent T cells transform to blasts. With continuing lymphocyte activation, progressively more astrocytes begin to release vesiculated material. This is followed by disruption of the membrane and complete lysis of the cells. After 30 h of contact with T cells and antigen, the originally contiguous astrocyte monolayers are almost completely disrupted (Fig. 1b).

We quantified and analysed T-cell-dependent destruction of antigen-presenting astrocytes using a conventional ⁵¹Cr-release assay. Examination of the T-cell dose response of astrocyte lysis after 6 h reveals an effector: target ratio of 1:1 (Fig. 2a). Furthermore, whereas the encephalitogenic Lewis T line Z1a released more than 50% of the astrocyte-incorporated isotope in the presence of MBP, a T-cell line specific for PPD (purified protein derivative of tuberculin) produced no detectable lysis (Fig. 2a). Astrocyte lysis by T cells depends strictly on the presence of MBP (Fig. 2b). Antigen pulse experiments indicate that the antigen seems to be taken up by the astrocytes (Table 1) and that a contact period of 1 h is sufficient for optimal subsequent presentation of MBP. Previous experiments have established that cell line Z1a, like all other permanent rat T-cell lines, expresses the membrane markers W3/13 and W3/25, but not OX8, and that such cells recognize their antigen only in the context of syngeneic Ia antigens^{7,8}. Ia determinants are also the critical restriction elements in astrocyte lysis. The monoclonal

Table 1 MBP uptake and presentation by Ia-induced astrocytes

T-cell line	MBP pulse		⁵¹ Cr release (%)
	Concentration (µg ml ⁻¹)	Duration (h)	
Z1a	200	18	59.9 ± 4.2
Z1a	200	6	60.4 ± 3.1
Z1a	200	3	63.0 ± 4.1
Z1a	20	3	56.8 ± 3.2
Z1a	200	1	60.0 ± 0.5
Z1a	0	—	24.2 ± 1.4
None	0	—	20.9 ± 0.6

Culture and assay conditions are described in Fig. 1 legend. Ia-induced astrocytes were incubated with varying MBP concentrations. After graded incubation periods, the antigen was removed from the cultures by twofold washing with medium. MBP-specific T cells (4×10^5) were added and isotope release was determined after another 16 h. The values represent the mean of duplicate experiments ($\pm$ s.d.).

anti-Ia antibody OX3, at a dilution of 1:1,000, suppressed lysis almost completely when measured after either 9 or 18 h of incubation (Table 2). Furthermore, astrocyte lysis is by no means limited to the Z1a line. Another Lewis-derived MBP-specific encephalitogenic T-cell line, L.BP 644, which had been selected from primed lymph node cells 2 months before this experiment, showed similar lytic capacity (Table 2). In striking contrast, T cells that recognize MBP epitopes distinct from the main encephalitogenic peptide sequence 68–88 and which are unable to mediate EAE, such as lines bs84 and L.C2, caused only marginal isotope release. Moreover, line L.OA, which is a Lewis line specific for ovalbumin, had no effect on astrocyte integrity, irrespective of the presence of MBP or ovalbumin.

We next investigated whether other types of accessory cells besides antigen-presenting astrocytes were susceptible to antigen-dependent cytotoxicity. We used Lewis rat peritoneal macrophages as antigen-presenting targets. These cells had been induced by intraperitoneal (i.p.) proteose-peptone injection and, like the astrocyte cultures, had been precultured with IFN- γ -containing conditioned media. Like astrocytes, Ia⁺ macrophages were lysed in the presence of antigen by encephalitogenic, but not by non-encephalitogenic MBP-specific T cells. In contrast, skin-derived fibroblasts were unable to present antigen and were resistant to T-cell lysis (Table 3).

Our results are surprising in several respects. First, the very fact that our MBP-specific T cells are cytolytic was unexpected. As mentioned above, all these lines are W3/25⁺ OX8⁺, a phenotype considered to define T 'helper/inducer' cells⁹. It should be mentioned that at least some of our encephalitogenic lines tested here were found to act as T-helper cells for syngeneic B cells producing specific antibodies¹⁰, and that earlier work was unable to detect cytolytic activity¹¹ when macrophages were

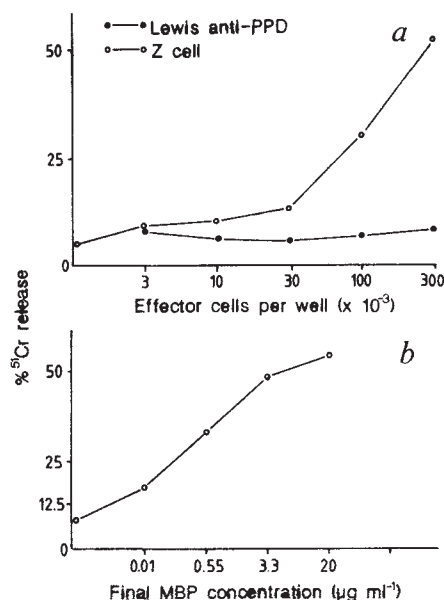


Fig. 2 Antigen-dependent cytotoxicity of encephalitogenic T cells (Z1a) against antigen-presenting syngeneic astrocytes. *a*, Dose-response of T-cell lines to astrocyte lysis; *b*, effect of MBP concentration on astrocyte lysis.

Methods. Ia-induced astrocytes were derived as for Fig. 1; 5×10^4 astrocytes were seeded in 0.5 ml of Complete Medium¹ supplemented with 10% fetal calf serum and 10% Con A supernatant into 48-well cluster plates (Costar). After 5 days, the cultures were labelled with 5 µCi Na⁵¹CrO₄ (Amersham) for 20 h. After three rinses, freshly activated T cells were added (4×10^5 ; *b*; graded doses, *a*), plus MBP (final concentration 10 µg ml⁻¹, *a*). After 9 h, the plates were centrifuged, the supernatants were removed and the sediments (including adherent cells) were lysed with 0.1 M NaOH. The radioactivity of supernatant and cell lysates was determined in parallel in a γ -spectrometer (Kontron). All experiments were performed in triplicate.

used as target cells. These studies did not show, however, whether Ia antigens were expressed on the cell surface. On the other hand, several human^{12–15} and murine^{16,17} T-cell lines have been reported to lyse target cells, although the effector cells expressed T4 and L3T4 phenotypes, respectively. Many data have demonstrated unequivocally that not all cytotoxic T cells are restricted by class I MHC determinants but that class II-restricted cytotoxicity may be important in normal anti-viral immune reactivity^{18,19}. However, it has been suggested that the cytolytic capacity of Ia-restricted T cells reflects a phenotypic switch induced by interleukin-2-like factors added to the culture media²⁰. The fact that a very 'young' T-cell line (L.BP644) was as effective as line Z1a, which has been kept in culture con-

Table 2 Cytolytic activity of antigen-specific T cells

T-cell line	Target antigen ³²	Antigen (10 µg ml ⁻¹)	Blocking antibody ^{33,34} (dilution 1:4,000)	⁵¹ Cr release (%)	
				Expt 1 (18 h)	Expt 2 (9 h)
Z1a	(MBP, 68–88)	MBP	0	80.5 ± 4.6	78.4 ± 6.6
		Ovalbumin	0	13.8 ± 0.1	4.4 ± 0.1
		MBP	OX3 (class II)	19.4 ± 3.2	15.7 ± 1.8
		MBP	OX18 (class I)	85.3 ± 2.5	64.5 ± 1.9
L.BP644	(MBP, 68–88)	—	0	15.0 ± 2.2	
		MBP	0	73.4 ± 1.6	
bs84	(MBP, undetermined)	MBP	0	33.6 ± 9.1	15.0 ± 1.0
L.C2	(MBP, 48–67)	MBP	0	39.1 ± 8.6	29.8 ± 2.8
L.OA	(Ovalbumin)	MBP	0	22.4 ± 8.6	5.2 ± 0.5
		Ovalbumin	0	21.6 ± 2.0	9.1 ± 0.3
None (spontaneous release)		—	0	13.2 ± 1.11	7.4 ± 0.2

Technical procedures are described in Fig. 2 legend.

Table 3 Cytotoxic effect of MBP-specific T cells on protease-peptone-induced rat peritoneal macrophages and skin-derived fibroblasts

T-cell line (4×10^5 per well)	Target	Antigen ($10 \mu\text{g ml}^{-1}$)	^{51}Cr release (%)
None	Mph	—	24.0 ± 1.2
Z1a	Mph	—	28.4 ± 1.8
Z1a	Mph	MBP	81.6 ± 5.6
bs84	Mph	—	27.3 ± 0.1
bs84	Mph	MBP	28.2 ± 2.7
None	Fibroblasts	—	17.0 ± 1.4
Z1a	Fibroblasts	—	18.4 ± 0.8
Z1a	Fibroblasts	MBP	20.7 ± 2.5

Lewis rats were injected i.p. with 15 ml of 10% thioglycollate solution (Serva). After 4 days, the peritoneal cavities were rinsed and macrophages isolated from the eluates. Fibroblasts were obtained from the skin of neonatal Lewis rats by trypsinization and used after three passages *in vitro*. Fibroblasts and macrophages (Mph) (4×10^4 cells per well) were cultured for 1 day in the presence of IFN- γ -containing conditioned media (mouse spleen cells cultured for 24 h with concanavalin A (Con A)), labelled with ^{51}Cr and exposed for 9 h to T cells.

tinuously for more than 4 years, argues against an *in vitro* 'ageing' effect in our system (Table 2).

The second unexpected observation is the apparent line-dependence of the killing effect. Out of a total of 12 antigen-specific T-cell lines and clones screened so far, none of the lines recognizing non-myelin antigens, such as ovalbumin or PPD, were lytic. Even recognition of different epitopes on the MBP molecule was associated with different intensities of target lysis by our T-cell-line panel. Highest lytic effects were seen with T cells recognizing the encephalitogenic peptide sequence (amino acids 68–88); T cells recognizing the adjacent, non-encephalitogenic sequence 48–67 (for example L.C2, Table 2) were much less cytotoxic to antigen-presenting astrocytes. Furthermore, considering the sequence homologies between viral and myelin proteins²¹, it is tantalizing to speculate that the nature of the epitope (co-)determines whether or not a T lymphocyte acts as a killer cell.

Finally, the fact that the cytolytic capacity of a MBP-specific T-cell line seems to correlate well with its encephalitogenic capacity deserves particular mention. Since astrocytes are the main Ia-inducible, potentially antigen-presenting cell population of the brain parenchyme²², could a cytotoxic interaction between encephalitogenic T cells and local astrocytes be involved in creating the functional EAE lesion? At present, local oedema formation²³ and secretion of blocking factors by effector T cells²⁴ or recruited macrophages^{25,26} are considered to be the main mechanisms in EAE. Cytotoxic killing of myelin-forming brain cells, previously favoured by many as the central element of immune demyelination and CNS inflammation, has become less popular since transfer studies showed that disease is mediated by T cells of putative T-helper phenotype^{3,8}, and since myelin-forming oligodendrocytes fail to express both MBP²⁷ and inducible Ia determinants^{22,28}. Our observations suggest that pathogenic interactions between cytotoxic MBP-specific Ia-restricted T cells and Ia-inducible brain cells may have a role in the pathogenesis of EAE. Brain endothelial cells, which express Ia antigens on treatment with IFN- γ and present antigen to primed T lymphocytes²⁹, would be potential targets, as would pericytic cells, which are often Ia⁺ in EAE brains³⁰. Moreover, cytotoxic MBP-specific T-cell lines that attack astrocytes induced to express Ia together with adapted MBP, could have a role in the pathogenesis of EAE. Impairment of astrocyte function could have multiple effects on neuronal functions via disturbance of the ion equilibrium, via trophic defects and by deficient compartmentalization of the brain parenchyme³¹.

In conclusion, our finding that W3/25⁺, OX8⁺ rat T cells are capable of cytotoxic target lysis raises further doubt about an

absolute association between T-cell membrane phenotype and immune function. Finally, if epitope-dependent lysis of astrocytes by autoaggressive T cells is confirmed as being an essential step in autoimmune encephalopathy, this phenomenon could be used to develop a cytolytic assay system for human encephalitogens, which could be helpful in diagnosing multiple sclerosis.

This unit is supported by funds of the Hermann-and-Lilly-Schilling Stiftung. We thank Ms Petra Bourquin and Roswitha Röhrig for T lines and astrocytes, and Mrs Barbara Goebel for preparing the manuscript.

Received 13 August 1985; accepted 6 January 1986.

- Ben-Nun, A., Wekerle, H. & Cohen, I. R. *Eur. J. Immun.* **11**, 195–199 (1981).
- Wekerle, H. *Lab. Invest.* **51**, 199–205 (1984).
- Swanborg, R. H. *J. Immun.* **130**, 1503–1505 (1983).
- Linington, C. *et al. J. Immun.* **133**, 1946–1950 (1984).
- Izumo, S., Linington, C., Wekerle, H. & Meyermann, R. *Lab. Invest.* (in the press).
- Fontana, A., Fierz, W. & Wekerle, H. *Nature* **307**, 273–276 (1984).
- Fierz, W., Endler, B., Reske, K., Wekerle, H. & Fontana, A. *J. Immun.* **134**, 3785–3793 (1985).
- Ben-Nun, A. & Cohen, I. R. *J. Immun.* **129**, 303–308 (1982).
- Mason, D. W. *et al. Immun. Rev.* **74**, 57–82 (1983).
- Schluessener, H. J. thesis, Univ. Würzburg (1984).
- Holoshitz, J., Naparstek, Y., Ben-Nun, A., Marquardt, P. & Cohen, I. *Eur. J. Immun.* **14**, 729–733 (1984).
- Biddison, W. E., Rao, P. E., Talle, M. A., Goldstein, G. & Shaw, S. *J. exp. Med.* **156**, 1065–1076 (1982).
- Krensky, A. M., Reiss, C. S., Mier, J. M., Strominger, J. L. & Burakoff, S. J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2365–2369 (1982).
- Meuer, S. C., Schlossman, S. F. & Reinherz, E. L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4395–4399 (1982).
- Spits, J., Borst, J., Terhorst, C. & deVries, J. E. *J. Immun.* **129**, 1563–1579 (1982).
- Cerottini, J.-C. & MacDonald, H. R. *J. Immun.* **126**, 490–496 (1981).
- Tite, J. P. & Janeway, J. A. *Eur. J. Immun.* **14**, 878–886 (1984).
- Roller, G. (ed.) *Immun. Rev.* **74** (1983).
- Jacobson, S., Nepom, G. L., Richert, J. R., Biddison, W. G. & McFarland, H. J. *exp. Med.* **161**, 263–268 (1985).
- Fleischer, B. *Nature* **308**, 365–367 (1984).
- Jahnke, U., Fischer, E. H. & Alvord, E. C. *Science* **229**, 282–284 (1985).
- Wong, G. H. W., Bartlett, P. F., Clark-Lewis, I., Battye, F. & Schrader, J. W. *Nature* **310**, 688–691 (1984).
- Paterson, P. Y. *Fedn Proc.* **35**, 2428–2434 (1976).
- Yarom, Y. *et al. Nature* **303**, 246–248 (1983).
- Brosnan, C. F., Bornstein, M. B. & Bloom, B. R. *J. Immun.* **126**, 614–620 (1981).
- Cammer, W., Bloom, B. R., Norton, W. T. & Gordon, S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1554–1558 (1978).
- Omlin, F. X., Webster, H. de F., Palkovits, C. G. & Cohen, S. R. *J. Cell Biol.* **95**, 242–248 (1982).
- Lisak, R. P. *et al. Brain Res.* **289**, 285–292 (1984).
- McCarran, R. M., Kempinski, O., Spatz, M. & McFarlin, D. E. *J. Immun.* **134**, 3100–3103 (1985).
- Vass, K., Lassmann, H., Wekerle, H. & Wisniewski, H. *Ann. Neuropath.* (in the press).
- Duffy, P. E. *Astrocytes Normal, Reactive and Neoplastic* (Raven, New York 1983).
- Chou, C.-H. J., Chou, F. C.-H., Kowalski, T. J., Shapira, R. & Kibler, R. F. *J. Neurochem.* **28**, 115–119 (1977).
- McMaster, W. R. & Williams, A. F. *Eur. J. Immun.* **9**, 426–433 (1979).
- Fukumoto, T., McMaster, W. R. & Williams, A. F. *Eur. J. Immun.* **12**, 237–243 (1982).

Unexpected expression of a unique mixed-isotype class II MHC molecule by transfected L-cells

Ronald N. Germain & Helen Quill

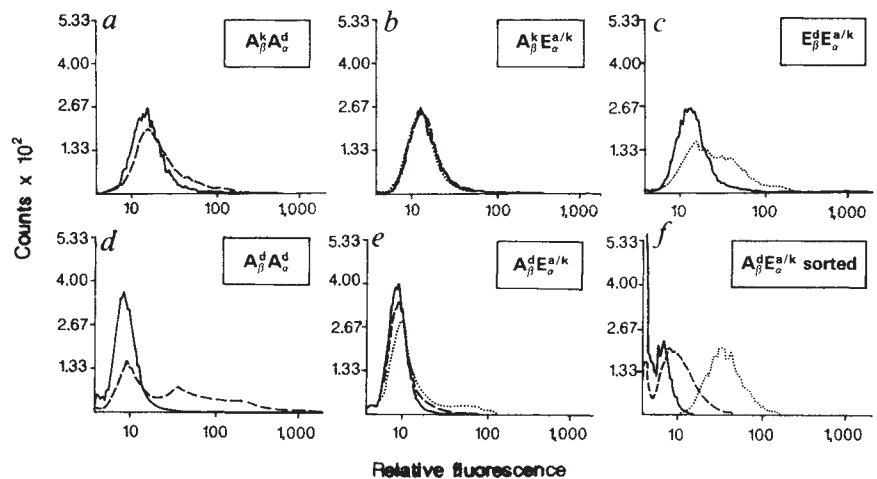
Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

Class II (Ia) major histocompatibility complex (MHC) molecules are heterodimeric integral membrane proteins composed of non-covalently linked α and β glycoprotein chains^{1,2}. Studies of both normal cells and L-cell transfectants have shown that neither α -nor β -chains are found on the cell surface alone, and that $\alpha\beta$ dimers are required for membrane expression^{1–6}. In both mouse and man, several distinct non-allelic α and β genes exist. Analysis of Ia molecules by immunoprecipitation and two-dimensional gel electrophoresis has demonstrated apparently selective association

Fig. 1 Surface expression of both A_β^d and $E_\alpha^{a/k}$ following co-transfection of L-cells. *a*, $A_\beta^k A_\alpha^d$ transfectants stained with 10.2.16²⁵ (A_β^k -specific monoclonal antibody). *b*, $A_\beta^k E_\alpha^{a/k}$ transfectants stained with 14.4.4S²⁶ (E_α -dependent monoclonal antibody) or with 10.2.16. *c*, $E_\beta^d E_\alpha^{a/k}$ transfectants stained with 14.4.4S or MKD6²⁷ (A_β^d -dependent monoclonal antibody). *d*, $A_\beta^d A_\alpha^d$ transfectants stained with MKD6. *e*, $A_\beta^d E_\alpha^{a/k}$ transfectants stained with 14.4.4S or MKD6. *f*, $A_\beta^d E_\alpha^{a/k}$ transfectant line after preparative cell sorting, stained with 14.4.4S or MKD6. Control staining (—) represents fluorescein isothiocyanate (FITC)-second antibody alone; ---, staining with anti-I- A_β antibody, 10.2.16 (*a*, *b*), or MKD6 (*d*–*f*), followed by FITC-anti-mouse immunoglobulin; ···, staining with anti-I- E_α antibody, 14.4.4S, followed by FITC-anti-mouse immunoglobulin.

Methods. Thymidine kinase-negative L-cells

(DAP.3) were co-transfected, using the calcium phosphate precipitation technique, with 5 or 10 μ g each of carrier-free α and β plasmid DNA, together with 1 ng of plasmid DNA containing the herpes simplex *tk* gene^{6,17}. The α - and β class II gene constructs used were: A_β^k = pgI- A_β^k -gpt-1 (ref. 18); A_β^d = pgI- A_β^d -gpt-49 (ref. 6); A_α^d = an A_α complementary DNA in the pcEXV expression vector^{14,19,20}; $E_\alpha^{a/k}$ = an 8-kilobase *Hind*III fragment containing an E_α gene from the (H -2^a $\times$ H -2^k) T-cell hybridoma 2B4, identified in a J11 phage library, and subcloned into the *Hind*III site of pUC8 (the designation *a/k* indicates uncertainty about the haplotype origin of this clone). Transfectants were selected in hypoxanthine-aminopterin-thymidine medium⁶, and pools of primary transfectants were stained with monoclonal anti-Ia antibody and fluorescent second antibody as described previously^{6,14}. The figure shows the results of analysing these stained cells on an EPICS V flow microcytometer, expressed as cell number (ordinate) versus relative fluorescence. *f*, The same analysis performed on cells prepared from the primary cells in *e* after preparative cell sorting for brightly staining cells.



of particular pairs of the various α - and β -chains to form the expressed class II isotypes I-A and I-E (mouse) or DQ, DP and DR (human)¹. Because the various α - or β -chains encoded by distinct loci exist in many allelic forms within a species^{1,2,7,8}, such specific pairing suggests a special role for isotypically conserved regions of each chain in the association process. In attempting to localize such putative assembly-controlling regions using the technique of DNA-mediated gene transfer, various combinations of murine α and β genes were introduced into L-cells. Here we report the unexpected observation, following transfection, of mixed-isotype ($A_\beta^d E_\alpha^{a/k}$) molecules on the L-cell membrane and document that the formation of this pair is strongly influenced by allelic polymorphism of the A_β chain.

The amino-acid sequences of allelic forms of A_α , A_β and E_β show the most striking conservation (that is, they display isotype-, rather than allele-specific sequences) in the β 2 or α 2, transmembrane and intracytoplasmic regions¹. This suggests that control of isotype selection during heterodimer assembly maps to these portions of α and β , in contrast to the important role recently demonstrated for the NH₂-terminal polymorphic domain of A_β in controlling intra-isotypic I-A expression⁹. To investigate this issue directly, L-cells were co-transfected with isotype-mismatched gene pairs ($A_\beta E_\alpha$). Based on previous immunoprecipitation results^{10,11}, no surface Ia expression was expected with this pair of chains, and thus, recombinant proteins derived from 'exon-shuffled' $A_\beta \leftrightarrow E_\beta$ or $A_\alpha \leftrightarrow E_\alpha$ genes could then be used to identify the regions involved in regulating isotype-selective chain pairing. Figure 1 shows the unexpected results of this experiment. Although no surface staining of primary transfectants receiving A_β^k plus $E_\alpha^{a/k}$ was seen with either monoclonal antibody 10.2.16 (specific for the β 1 domain of A_β^k) or monoclonal antibody 14.4.4S (which detects an epitope associated with $E_\alpha^{a/k}$ chains) (Fig. 1b), cells receiving A_β^d and $E_\alpha^{a/k}$ showed low but significant reactivity with both monoclonal antibody MKD6 (dependent on A_β^d for the formation of the relevant epitope) and 14.4.4S (Fig. 1e). The degree of staining with 14.4.4S was as great as that achieved with 10.2.16 for cells receiving the isotype-matched, but haplotype-mismatched $A_\beta^k A_\alpha^d$ pair (Fig. 1a). The isotype-matched $E_\beta^d E_\alpha^{a/k}$ (Fig. 1c) and $A_\beta^d A_\alpha^d$ (Fig. 1d) controls both showed a greater number of strongly positive cells. Cells transfected with A_β^d alone or $E_\alpha^{a/k}$ alone

showed no staining above background with MKD6 or 14.4.4S, respectively (results not shown).

Figure 2 shows the fluorescent profiles of one line generated by repetitive cell sorting of $A_\beta^d E_\alpha^{a/k}$ transfectants and one line created using $E_\beta^d E_\alpha^{a/k}$. Both lines stained strongly with 14.4.4S. The line prepared using A_β^d stained with MKD6, but less well than it stained with 14.4.4S. The level of staining with MKD6 varied with different preparations of this antibody—the results shown illustrate the weakest staining seen. However, in all cases the staining was lower than that observed with 14.4.4S, suggesting a decreased affinity of MKD6 for A_β^d in association with E_α rather than its normal partner A_α^d . MKD6 did not stain the $E_\beta^d E_\alpha^{a/k}$ transfectant at all. Northern blot analysis confirmed the presence of A_β , but not E_β , transcripts in the $A_\beta^d E_\alpha^{a/k}$ cells, and of E_β , but not A_β , transcripts in the $E_\beta^d E_\alpha^{a/k}$ cells (Fig. 2).

To determine whether isotype-mismatched heterodimers were actually formed by the $A_\beta^d E_\alpha^{a/k}$ transfectant, or whether the presence of α - and β -chains in the cell somehow facilitated membrane expression of independent chains, immunoprecipitation and two-dimensional gel electrophoretic analyses were performed on lysates from internally labelled aliquots of the sorted cells shown in Fig. 2a and from two control transfectants bearing $A_\beta^d A_\alpha^d$ or $E_\beta^d E_\alpha^{a/k}$ molecules. Control gels showed that 14.4.4S precipitated class II molecules only from the I-E-, but not I-A-expressing control cells, while the monoclonal antibody 34.5.3 (anti-I- A_β^d) and a rabbit anti-I-A antiserum precipitated class II molecules only from the I-A-expressing, but not I-E-bearing cells (not shown). Figure 3 shows the autoradiographs of two-dimensional gel electrophoretic analysis of immunoprecipitates from the $A_\beta^d E_\alpha^{a/k}$ transfectant. All three antibodies (14.4.4S, 34.5.3 and rabbit anti-I-A) precipitated class II molecules from this cell, and showed spots characteristic of E_α and A_β . The anti-I-E antibody precipitated a clearly disproportionate ratio of E_α to A_β , suggesting the presence of E_α chains recognizable by the antibody which were not assembled with A_β . The anti-I-A antibodies precipitated a more equal ratio of A_β and E_α . The co-precipitation of both A_β and E_α chains by a single antibody directed at only one or the other chain provides direct evidence for the existence of $A_\beta E_\alpha$ heterodimers in these transfectants.

These results are in marked contrast to those of many previous

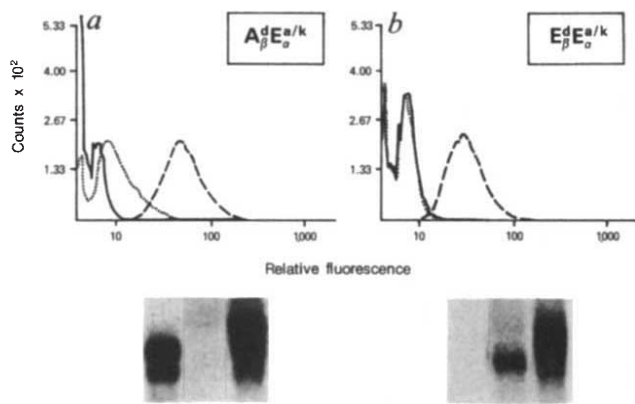


Fig. 2 Flow microfluorimetric and Northern blot analysis of immunoselected transfectants expressing $A_{\beta}^d E_{\alpha}^k/k$. L-cells transfected with $E_{\beta}^d E_{\alpha}^k/k$ or $A_{\beta}^d E_{\alpha}^k/k$ were repetitively sorted on the EPICS V flow microfluorimeter, and lines showing uniform moderate levels of surface Ia expression derived. Such cells were stained with MKD6 (·····) or 14.4.4S (----) or were used to prepare total cytoplasmic RNA for Northern blot analysis using A_{β} , E_{β} or E_{α} -specific cDNA probes^{9,21}.

studies on Ia molecules. However they are consistent with our recent demonstration of $A_{\beta}DR_{\alpha}$ pairing in transfected human Epstein-Barr virus (EBV)-transformed lymphoblastoid cells¹². Further, Lerner *et al.*¹³ noted the presence of autoradiographic spots 'indistinguishable' from A_{α} in immunoprecipitates made using the anti-I-E monoclonal antibody Y-17. These putative A_{α} spots were seen in the absence of A_{β} , but in the presence of E_{β} and E_{α} , and were thought possibly to reflect $E_{\beta}A_{\alpha}$ molecules recognized by Y17. However, no direct proof of the existence of such molecules has been presented since the initial observation. If these data are correct, it would be expected that $E_{\beta}A_{\alpha}$ transfectants would show Y17 surface reactivity, and this is currently being examined.

One observation that is somewhat inconsistent with the results of our L-cell transfections is the report of surface expression of E_{α} alone by A.TFR5 cells, which fail to synthesize E_{β} chains¹¹. To date, no L-cell transfected with individual A_{β} , A_{α} , E_{β} or E_{α} genes or the human equivalents has shown surface expression of the respective protein product^{3-6,14}, even though monoclonal antibodies or heteroantisera potentially capable of binding such molecules were used. Although additional experiments with class II genes covalently linked to drug selection markers, and under strong promoter/enhancer control, need to be done to verify these negative results, one explanation of the apparent discrepancy is that A.TFR5 in fact resembles the $A_{\beta}E_{\alpha}$ transfectants reported here in that it synthesizes A_{β} and E_{α} without E_{β} . The low concentration of E_{α} detected on the surface membrane may thus represent $A_{\beta}E_{\alpha}$ pairs expressed at a low (but detectable) level, rather than E_{α} expressed by itself.

Although the above examples are consistent with the present observations, an explanation for the scarcity of reports on such mixed isotype molecules is still needed. One possibility is that even though $A_{\beta}^d E_{\alpha}^k/k$ molecules were found under the present experimental conditions, isotypically conserved portions of α - and β -chains still play an important part in favouring intra-isotypic pairing. In cells expressing a full set of α - and β -chains, competition would favour assembly of the preferred intra-isotype pairs, giving few or no mixed molecules. The use of inter-isotype recombinant α - or β -chains, and direct testing of competition between A_{β} and E_{β} for E_{α} in a single transfected cell, will provide additional insight into this issue.

Although L-cells expressing $A_{\beta}^d E_{\alpha}^k/k$ were readily detected, surface expression of $A_{\beta}^d E_{\alpha}^k/k$ was not observed. This influence of A_{β} allelic polymorphism on assembly/expression with E_{α} is

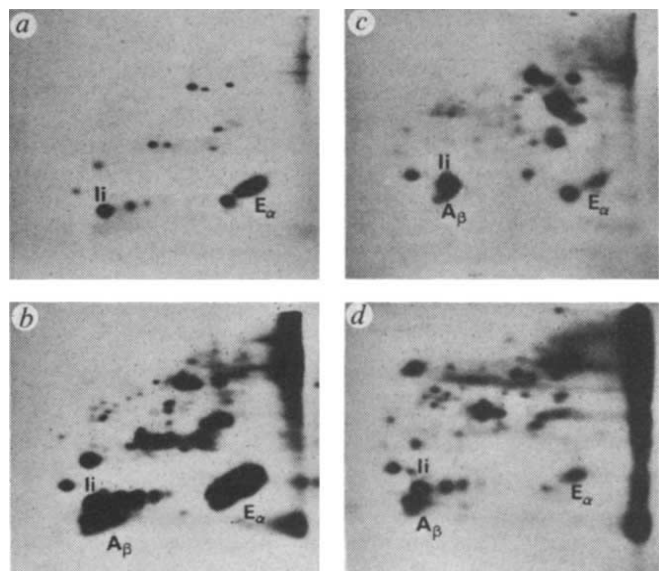


Fig. 3 Assembly of $A_{\beta}^d E_{\alpha}^k/k$ heterodimers. Cells from the immunoselected $A_{\beta}^d E_{\alpha}^k/k$ transfectant shown in Fig. 2a, were labelled with a combination of ³⁵S-cysteine and methionine, and cell lysates were subjected to immunoprecipitation and two-dimensional gel electrophoresis, using non-equilibrium pH gradient electrophoresis (NEPHGE) in the first dimension, and SDS-polyacrylamide gel electrophoresis in the second²². The autoradiographs show the results obtained using monoclonal antibody 14.4.4S (a, b), monoclonal antibody 34.5.3 (anti-I- A_{β}^d)²³ (c) and rabbit anti-I-A heteroantiserum²⁴ (d) for precipitation. Short (a) and long (b) exposures for the 14.4.4S gels were for 6 and 26 days, respectively. li, invariant chain; A_{β} , A_{β}^d polypeptide; E_{α} , E_{α}^k/k polypeptide. The pH gradient runs from left (basic) to right (acidic).

in agreement with previous studies from this laboratory on $A_{\beta}A_{\alpha}$ expression⁹. This previous work demonstrated striking variations in the concentration of Ia expression following co-transfection of A_{α} together with A_{β} from a different haplotype. Studies using 'exon-shuffled' A_{β} genes revealed that the polymorphic $\beta 1$ domain controlled this effect and preliminary results suggest that this same region regulates the differential $A_{\beta}E_{\alpha}$ pairing observed here (A. Sant, N. Braunstein and R.N.G., unpublished observations). The demonstration of a similar effect of allelic polymorphism on assembly with both A_{α} and E_{α} suggests a conserved role for the amino-terminal region of the β -chain in Ia assembly, irrespective of additional 'framework' or isotypic influences on this process.

The documentation of inter-isotypic pairing may have significance for studies of Ia restriction and *Ir* gene function. The present results suggest that inter-isotypic pairing contributes to the diversity of Ia molecules potentially available to serve as restriction elements for T-helper cell recognition of antigen. The E_{α} -null haplotypes *b* and *s*¹⁵ have long been used to map the control of immune phenomena to the I-E molecule¹⁶. In fact, given the results reported here, such experiments only map the response with respect to E_{α} ; $E_{\beta}A_{\alpha}$ molecules may have a previously unsuspected role in such situations. Similarly, antibody blocking studies with anti-I-A reagents which are actually anti- A_{β} may falsely map responses to $A_{\beta}A_{\alpha}$ molecules, when in fact $A_{\beta}E_{\alpha}$ pairs are crucial. The possible low-level expression of such molecules may lead to a greater than normal dependence on adjuvant effects in *Ir* systems, as such agents may act largely by evoking lymphokine-mediated increases in the level of Ia. Experiments are in progress to test directly for the existence of T cells restricted to $A_{\beta}E_{\alpha}$ class II molecules.

The existence of mixed-isotype molecules whose expression depends critically on the particular alleles involved may be

relevant to HLA-DR-linked diseases. These molecules may provide a 'responder' *Ir* gene phenotype for certain environmental antigens, or even for self-antigens; this would account for haplotype- rather than locus-specific disease associations. Direct demonstration of isotypically mixed human Ia molecules, and examination of the possible correlation of their expression with disease may prove particularly interesting.

We thank Drs Ronald Schwartz, Andrea Sant and Ned Braunstein for comments on the manuscript, Alessio Palini for performing the flow cytometry, and Shirley Starnes for editorial assistance.

Received 16 September; accepted 15 November 1985.

1. Kaufman, J. F., Affray, C., Korman, A. J., Shackelford, D. A. & Strominger, J. *Cell* **36**, 1-13 (1984).
2. Hood, L., Steinmetz, M. & Malissen, B. A. *Rev. Immun.* **1**, 529-568 (1983).
3. Rabourdin-Combe, C. & Mach, B. *Nature* **303**, 670-674 (1983).
4. Malissen, B., Steinmetz, M., McMillan, M., Pierres, M. & Hood, L. *Nature* **305**, 440-443 (1983).
5. Malissen, B. *et al. Cell* **36**, 319-327 (1984).
6. Norcross, M. A., Bentley, D. M., Margulies, D. H. & Germain, R. N. *J. exp. Med.* **160**, 1316-1337 (1984).
7. Klein, J., Figueroa, F. & David, C. S. *Immunogenetics* **17**, 553-596 (1983).
8. Bukara, M., Vincek, V., Figueroa, F. & Klein, J. *Immunogenetics* **21**, 569-579 (1985).
9. Germain, R. N., Bentley, D. B. & Quill, H. *Cell* **43**, 233-242 (1985).
10. Jones, P. P., Murphy, D. B. & McDevitt, H. O. *J. exp. Med.* **148**, 925-939 (1978).
11. Murphy, D. B., Jones, P. P., Loken, M. R. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5404-5408 (1980).
12. Norcross, M. *et al.* (submitted).
13. Lerner, E. A. *et al. J. exp. Med.* **152**, 1085-1101 (1980).
14. Germain, R. N. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 2940-2944 (1985).
15. Mathis, D., Benoist, C., Williams, V., Kanter, M. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **40**, 273-277 (1983).
16. Schwartz, R. H. A. *Rev. Immun.* **3**, 237-262 (1985).
17. Margulies, D. H. *et al. J. Immun.* **130**, 463-470 (1983).
18. Germain, R. N., Norcross, M. A. & Margulies, D. H. *Nature* **306**, 190-194 (1983).
19. Davis, M. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 2194-2198 (1984).
20. Miller, J. *et al. J. Immun.* **134**, 4212-4217 (1985).
21. Rogers, M. J., Germain, R. N., Hare, J., Long, E. & Singer, D. S. *J. Immun.* **134**, 630-636 (1985).
22. O'Farrell, P. Z., Goodman, H. M. & O'Farrell, P. H. *Cell* **12**, 1133-1142 (1977).
23. Ozato, K., Mayer, N. & Sachs, D. H. *Transplantation* **34**, 113-120 (1982).
24. Robinson, R. R., Germain, R. N., McKean, D. J., Mescher, M. & Seidman, J. G. *J. Immun.* **131**, 2025-2031 (1983).
25. Oi, V. T., Jones, P. P., Goding, J. W., Herzenberg, L. A. & Herzenberg, L. A. *Curr. Top. Microbiol. Immun.* **7**, 115-129 (1978).
26. Ozato, K., Mayer, N. & Sachs, D. H. *J. Immun.* **124**, 533-540 (1980).
27. Kappler, J., Skidmore, B., White, J. & Marrack, P. *J. exp. med.* **153**, 1198-1214 (1981).

Expression of functional human interleukin-2 receptor in mouse T cells by cDNA transfection

Shigeru Kondo, Akira Shimizu, Michiyuki Maeda*, Yutaka Tagaya†, Junji Yodoi† & Tasuku Honjo

Department of Medical Chemistry, and † Institute for Immunology, Faculty of Medicine, and * Chest Diseases Research Institute, Kyoto University, Sakyo-ku, Kyoto 606, Japan

Interleukin-2 (IL-2) in combination with the IL-2 receptor has an essential role in antigen-stimulated proliferation of T lymphocytes¹⁻³. It has been proposed that the constitutive expression of the IL-2 receptor on adult T-cell leukaemia (ATL) cells⁴ may be associated with transformation of T cells^{5,6}. Although we and others⁷⁻⁹ have isolated complementary DNA clones encoding a protein that binds IL-2, formal proof that this protein is the IL-2 receptor requires demonstration of IL-2-dependent growth stimulation of cells expressing the protein. In addition, a functional assay system other than binding of IL-2 is required to investigate the molecular mechanism of signal transmission through the IL-2 receptor using artificially mutated cDNA. The IL-2 receptor expressed in non-lymphoid cells by cDNA transfection did not mediate a growth signal^{10,11}, implying that lymphoid cells expressing the functional receptor might have specific accessory molecule(s) for signal transmission by the receptor. Therefore, we established a line of IL-2-dependent mouse cells (CT/hR) express-

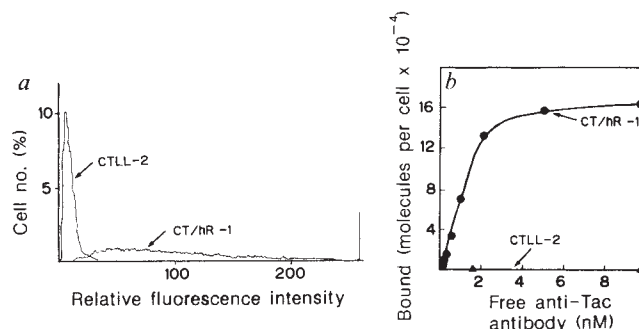


Fig. 1 *a*, Flow cytometric analysis of the human IL-2 receptor. CTLL-2 and CT/hR-1 were stained with the anti-Tac monoclonal antibody as described below. The result for CT/hR-2 (not shown) was almost identical to that obtained for CT/hR-1. *b*, Binding of ¹²⁵I-labelled anti-Tac monoclonal antibody to mouse T cells. ●, CT/hR-1; ○, CTLL-2.

Methods. Transfection of CTLL-2 was done according to a modified method of Queen and Baltimore²¹. CTLL-2 cells (1×10^7) were washed twice with serum-free Dulbecco's modified Eagle's medium (DMEM), and suspended in 1 ml of DEAE-DNA solution (400 µg DEAE-dextran, 3 µg pKCR-Tac2A, 1 µg pSV2-Neo in 1 ml DMEM). pKCR-Tac2A and pSV2-Neo were linearized with *Sal*I and *Bam*HI, respectively. The mixture was incubated at 37 °C for 20 min with occasional shaking. The cells were washed once with heparin solution (2 µg ml⁻¹ heparin in DMEM), and transferred in the culture medium (RPMI 1640 containing 10% fetal calf serum, 0.05 mM 2-mercaptoethanol and 10% rat spleen conditioned medium). Rat spleen conditioned medium (CM) containing rat IL-2 was prepared as described elsewhere¹². After 24 h the cells were transferred to the selection medium containing 0.4 mg ml⁻¹ of the antibiotic G418 (Gibco). Twenty neomycin-resistant clones were isolated and tested for expression of the human IL-2 receptor by staining with the anti-Tac monoclonal antibody. Two clones, CT/hR-1 and CT/hR-2, were stained and analysed in detail. CTLL-2 and CT/hR were washed with phosphate-buffered saline (PBS) and incubated in 20 µl of anti-Tac antibody solution (0.1 mg ml⁻¹ antibody in PBS) at 4 °C for 30 min, then the cells were washed and incubated in fluorescein isothiocyanate (FITC)-conjugated anti-mouse IgG antibody (Cappel) before analysis using a fluorescence-activated cell sorter (EPICS V, Coulter). The binding of ¹²⁵I-anti-Tac antibody was assayed as described previously²². Before assay, cells were incubated in IL-2-free medium for several hours to prevent contamination with cold IL-2 from the culture medium. Aliquots (200 µl) of 10⁶ cells each were incubated for 30 min at 37 °C in the presence of various amounts of anti-Tac antibody. Free and cell-bound antibodies were separated by centrifugation through an olive oil cushion and measured using a γ-counter. The extent of nonspecific binding was measured by incubating the cells in the presence of a 100-fold excess of unlabeled antibody, and this value was subtracted from the bound counts. ¹²⁵I-labelling of the anti-Tac antibody was done according to the Bolton-Hunter method²³. Labelling reagents were from NEN. The specific activity and relative molecular mass (*M_r*) of labelled anti-Tac antibody were 5,600 c.p.m. ng⁻¹ and 170,000, respectively.

ing both murine (endogenous) and human IL-2 receptors. Here, by blocking the endogenous mouse IL-2 receptors with monoclonal antibodies, we show that the human IL-2 receptor of CT/hR cells is functionally active. Although CT/hR expressed the human IL-2 receptor constitutively, growth of these cells was strictly dependent on IL-2, indicating that uncontrolled over-expression of the IL-2 receptor was not by itself sufficient for T-cell transformation.

A mouse cytotoxic T-cell line (CTLL-2)¹² which can grow only in the presence of IL-2 was transfected with plasmid pKCR-Tac2A containing human IL-2 receptor cDNA⁷. The promoter of the pKCR vector is derived from simian virus 40, which is highly active in many cultured cells. The plasmid pSV2-Neo was co-transfected as a selection marker. Two clones (CT/hR-1 and CT/hR-2) among 20 neomycin-resistant clones expressed human IL-2 receptors on their surface. Figure 1a shows a flow cytometric analysis of CT/hR-1 and CTLL-2 using a monoclonal antibody against the human IL-2 receptor (anti-

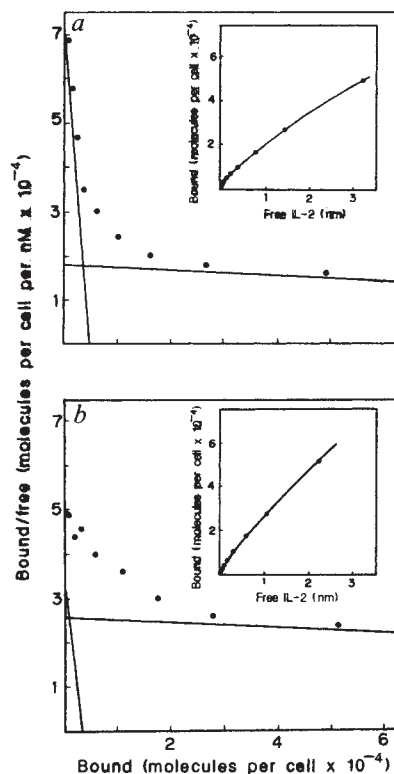


Fig. 2 Scatchard plot analyses of equilibrium binding of human recombinant ^{125}I -IL-2 to CTLL-2 (a) and CT/hR (b). Insets show normal plots of binding kinetics. The binding of ^{125}I -labelled IL-2 was assayed as described in Fig. 1 legend. The specific activity and M_r of ^{125}I -labelled IL-2 were 8,000 c.p.m. ng^{-1} and 15,000, respectively. Numbers and dissociation constants (K_d) of the high-affinity sites were determined using a computer-assisted geometrical simulation in which two asymptotic lines were calculated assuming that the Scatchard plot curves of CTLL-2 and CT/hR-1 were hyperboles. Nonspecific binding of ^{125}I -IL-2 to CTLL-2 and CT/hR was 0.71% and 0.60% of input counts, respectively, as measured by the addition of a 200-fold excess of unlabelled IL-2.

Tac¹³). More than 98% of CT/hR-1 cells were stained with the anti-Tac monoclonal antibody. The mean fluorescence intensity was approximately equal to that of MT-1, a human T-lymphotropic virus type I (HTLV-I)-transformed T-cell line bearing 5×10^5 IL-2 receptor molecules per cell¹¹. The parental CTLL-2 cells were not stained at all. The number of Tac antigens per CT/hR-1 cell (1.6×10^5) was estimated using ^{125}I -labelled anti-Tac antibody (see the saturation curve in Fig. 1b).

As human IL-2 binds to both human and murine IL-2 receptors, we studied the kinetics of binding of human recombinant IL-2 to CTLL-2 and CT/hR-1 cells (Fig. 2). CTLL-2 had 1.7×10^5 low-affinity sites ($K_d = 11 \text{ nM}$) and 5×10^3 high-affinity sites ($K_d = 71 \text{ pM}$) per cell. CT/hR-1 had 2.9×10^5 low-affinity sites ($K_d = 12 \text{ nM}$) and 3×10^3 high-affinity sites ($K_d = 87 \text{ pM}$) per cell. Although CT/hR-1 had 1.7 times more IL-2 binding sites than CTLL-2, CT/hR-1 had a similar number of the high-affinity sites. The results are consistent with the hypothesis that the high-affinity IL-2 receptor is a complex of the receptor and another molecule expressed in limited numbers specifically in the T cell. There exists some evidence that other molecules bind to the IL-2 receptor on the surface of T cells¹⁴.

To test the activity of the human IL-2 receptor on CT/hR-1, we blocked the activity of the endogenous murine IL-2 receptor using anti-mouse IL-2 receptor antibody (AMT13)¹⁵ which did not interfere with the human IL-2 receptor. The inhibition by this antibody of ^3H -thymidine uptake by CTLL-2 cells was tested in the presence of two concentrations of human IL-2 (Fig. 3a). At $2.5 \times 10^{-5} \mu\text{g ml}^{-1}$ of human IL-2, which gives a half-maximal

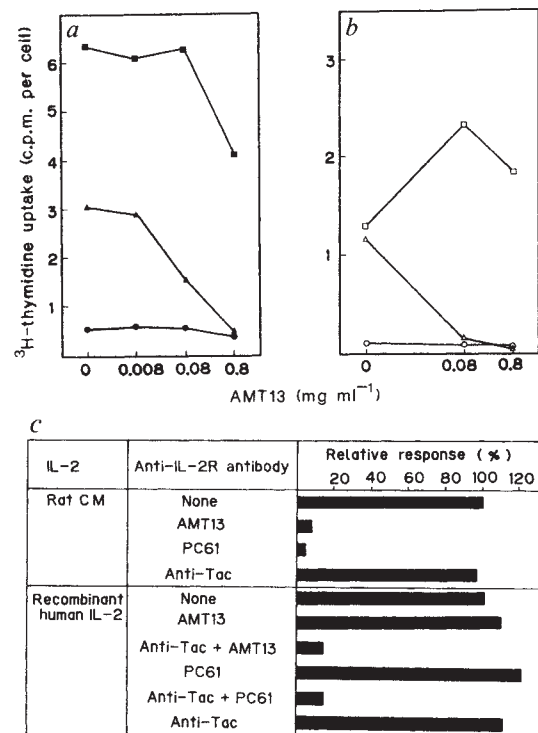


Fig. 3 a, Inhibition by AMT13 monoclonal antibody of ^3H -thymidine uptake by CTLL-2 cells. ●, Minus recombinant human IL-2; ■, ▲, plus 2.5×10^{-4} and $2.5 \times 10^{-5} \mu\text{g ml}^{-1}$ human recombinant IL-2, respectively. b, Absence of inhibition of human IL-2-dependent ^3H -thymidine uptake by CT/hR-1 cells treated with anti-mouse IL-2 receptor antibody (AMT13). CT/hR-1 cells were cultured in the presence of $2.5 \times 10^{-5} \mu\text{g ml}^{-1}$ human IL-2 (□) or 0.9 U ml^{-1} of IL-2-equivalent rat conditioned medium (Δ), or in the absence of the growth factor (○). c, Blocking of the human IL-2 receptor (IL-2R) activity of CT/hR-1 by anti-Tac antibody. Background incorporation without added IL-2 was subtracted from the values obtained, and the relative responses were expressed as per cent of ^3H -thymidine incorporation without antibody. The final concentrations of antibodies were as follows: AMT13, 0.8 mg ml^{-1} ; PC61, 10% of 100-fold-concentrated serum-free culture supernatants derived from the hybridoma (this amount of PC61 completely blocked the response of CTLL-2 to $1.75 \times 10^{-5} \mu\text{g ml}^{-1}$ of recombinant human IL-2); anti-Tac, 1 mg ml^{-1} , 0.6 U ml^{-1} of IL-2-equivalent rat CM and $1.75 \times 10^{-5} \mu\text{g ml}^{-1}$ of human recombinant IL-2 were used.

Methods. CTLL-2 or CT/hR-1 cells (1×10^4) were incubated for 4 h in 0.1 ml of growth factor-free medium (described in Fig. 1 legend) except that 0.5 mM 2-mercaptoethanol was used for the experiments with PC61. Then, 0.1 ml of medium containing the growth factor and anti-IL-2 receptor antibodies was added. After incubation at 37°C for 20 h, ^3H -thymidine (250 pCi in 50 μl of medium, 2 Ci mmol^{-1}) was added and the incubation continued for a further 4 h. Cells were collected using a microharvester and acid-insoluble radioactivities were counted using a liquid scintillation counter. Data are the mean of triplicate experiments.

response in CTLL-2, the antibody (0.8 mg ml^{-1}) completely inhibited the IL-2-dependent ^3H -thymidine uptake into CTLL-2. As the AMT13 antibody blocks the IL-2 receptor by competing with IL-2 (ref. 16) for the binding sites, 10-fold more antibody molecules were required to inhibit the growth of CTLL-2 at a 10-fold higher concentration of IL-2 ($2.5 \times 10^{-4} \mu\text{g ml}^{-1}$).

In the presence of $2.5 \times 10^{-5} \mu\text{g ml}^{-1}$ of human IL-2 or 0.9 unit ml^{-1} of rat IL-2 (supplied as spleen conditioned medium), CT/hR-1 incorporated about a half-maximal amount of ^3H -thymidine. The activity of 0.9 unit ml^{-1} of rat IL-2 was completely blocked by AMT13 ($\geq 0.08 \text{ mg ml}^{-1}$), in agreement with the fact that rat IL-2 cannot interact with the human IL-2 receptor². However, when rat IL-2 was replaced by human IL-2,

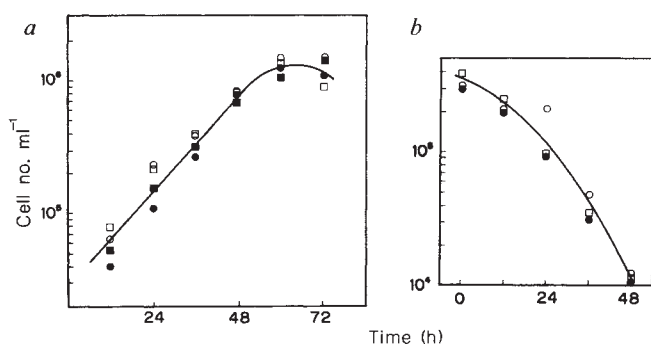


Fig. 4 Growth curves of CTLL-2 and CT/hR-1 in the presence (a) or absence (b) of IL-2. a, Cells were grown in either culture medium containing human recombinant IL-2 (0.05 $\mu\text{g ml}^{-1}$) or rat CM (10% or 12.5 U ml^{-1} IL-2 equivalent). b, Cells were washed three times with RPMI 1640 before starting the experiments and were grown in the culture medium without IL-2. Aliquots of cells were taken at the time intervals indicated and viable cells counted under the microscope after staining with Trypan blue. In a, symbols are: $\circ$, CTLL-2 with human IL-2; $\square$, CTLL-2 with rat CM; $\bullet$, CT/hR-1 with human IL-2; and $\blacksquare$, CT/hR-1 with rat CM. In b, symbols are: $\bullet$, CTLL-2; $\circ$, CT/hR-1; and $\square$, CT/hR-2.

the thymidine uptake by CT/hR-1 cells was not inhibited even in the presence of 0.8 mg ml^{-1} AMT13 (Fig. 3b), which blocked completely the effect of the same amount of human IL-2 on CTLL-2 (Fig. 3a). Another monoclonal antibody (PC61)¹⁷ against the mouse IL-2 receptor also blocked the CT/hR-1 response to rat IL-2 but did not block the response to human IL-2 (Fig. 3c). These results indicate that the human IL-2 receptor expressed on CT/hR-1 is functionally active and that the effective number of human IL-2 receptors on CT/hR-1 is more than required for the full response of CT/hR-1 to the limited amount of IL-2 added. As the anti-Tac antibody alone did not affect the growth of CT/hR-1, the number of murine IL-2 receptors is also abundant enough to support the full growth response. However, human IL-2-dependent thymidine uptake by CT/hR-1 was abolished completely by a simultaneous addition of the monoclonal antibodies against human (anti-Tac) and murine (AMT13 or PC61) IL-2 receptors. Taken together, these results show that the human IL-2 receptor expressed on CT/hR-1 is able to transmit a growth signal, in contrast to the receptor expressed on non-lymphoid cells^{10,11}.

The present study represents the first demonstration that an active receptor for a peptide hormone can be reconstituted by cDNA transfection and also represents the most direct evidence that our cDNA encodes the functional IL-2 receptor. We were unable to determine whether or not human IL-2 receptors of CT/hR cells contained high-affinity species, however, it is reasonable to assume this is the case because only the high-affinity receptor is active in signal transmission¹⁸. This and previous^{10,11} results indicate that some factor(s) absent in non-lymphoid cells is necessary for signal transmission through the IL-2 receptor. The murine factor should be able to act on both murine and human IL-2 receptors, and it may interact with conserved portions of the murine and human IL-2 receptors, such as the transmembrane and cytoplasmic regions¹⁹.

ATL cells bear a large number of IL-2 receptors when cultured *in vitro*, and most of them grow in the absence of IL-2 (refs 4-6). It was proposed that constitutive overexpression of the IL-2 receptor might be associated with the leukaemogenesis of ATL cells^{5,6}. The growth rate of CT/hR-1 was almost equal to that of CTLL-2 in the presence of a sufficient amount of human or rat IL-2 (Fig. 4a), suggesting that transfection of CTLL-2 with the human IL-2 receptor cDNA did not have a significant effect on the growth rate of this T cell. Removal of IL-2 from the culture medium brought about quick death of not only CTLL-2 but also CT/hR-1 and CT/hR-2 cells (Fig. 4b); in other

words, CT/hR-1 and CT/hR-2 remained strictly dependent on IL-2 although they expressed about the same number of the IL-2 receptors constitutively as ATL cell lines. Clearly, constitutive expression of the IL-2 receptor is not sufficient for autonomous growth of T cells. It is conceivable that overexpression of the growth factor receptor induced by HTLV-I infection is the first 'hit' in the process of leukaemogenesis²⁰, and that IL-2-independent growth of the T cell may require the second 'hit'.

We thank Dr T. Diamantstein for AMT13 antibody, Dr M. Nabholz for PC61 antibody, Dr T. Uchiyama for anti-Tac antibody, and Takeda Chemical Co. for recombinant IL-2. We also thank Ms A. Kasahara and Ms Y. Utsunomiya for technical assistance, and Ms K. Hirano for preparing the manuscript. This investigation was supported in part by grants for the special promotion of science, by a grant-in-aid from the Ministry of Education, Science and Culture of Japan, and by grants from the Mochida Foundation and the Yamanouchi Foundation.

Received 26 September 1985; accepted 14 January 1986.

1. Morgan, D. A., Ruscetti, F. W. & Gallo, R. C. *Science* **193**, 1007-1008 (1976).
2. Smith, K. A. *Immun. Rev.* **51**, 337-357 (1980).
3. Hemler, M. E., Brenner, M. B., McLean, J. M. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2172-2175 (1984).
4. Hattori, T., Uchiyama, T., Toibana, T., Takatsuki, K. & Uchino, H. *Blood* **58**, 645-647 (1981).
5. Tsudo, M., Uchiyama, T., Uchino, H. & Yodoi, J. *Blood* **61**, 1014-1016 (1983).
6. Yodoi, J., Uchiyama, T. & Maeda, M. *Blood* **62**, 509-510 (1983).
7. Nikaido, T. *et al.* *Nature* **311**, 631-635 (1984).
8. Leonard, W. J. *et al.* *Nature* **311**, 626-631 (1984).
9. Cosman, D. *et al.* *Nature* **312**, 768-776 (1984).
10. Greene, W. C. *et al.* *J. exp. Med.* **162**, 363-368 (1985).
11. Sabe, H. *et al.* *Molec. Biol. Med.* (in the press).
12. Gillis, S., Ferm, M. M., Ou, W. & Smith, K. A. *J. Immun.* **120**, 2027-2032 (1978).
13. Uchiyama, T., Broder, B. & Waldmann, T. A. *J. Immun.* **126**, 1393-1397 (1981).
14. von Boehmer, H., Crisanti, A., Kisielow, P. & Haas, W. *Nature* **314**, 539-540 (1985).
15. Osawa, H. & Diamantstein, T. *J. Immun.* **132**, 2445-2450 (1984).
16. Osawa, H. & Diamantstein, T. *Eur. J. Immun.* **15**, 299-301 (1985).
17. Ceredig, R., Lowenthal, J. W., Nabholz, M. & MacDonald, H. R. *Nature* **314**, 98-100 (1985).
18. Robb, R. J., Green, W. C. & Rusk, C. M. *J. exp. Med.* **160**, 1126-1146 (1984).
19. Shimizu, A. *et al.* *Nucleic Acids Res.* **13**, 1505-1516 (1985).
20. Maeda, M. *et al.* *J. exp. Med.* **162**, 1269-1274 (1985).
21. Queen, C. & Baltimore, D. *Cell* **33**, 741-748 (1983).
22. Robb, R. J., Munck, A. & Smith, K. A. *J. exp. Med.* **154**, 1455-1474 (1981).
23. Bolton, A. E. & Hunter, W. M. *Biochem. J.* **133**, 529-531 (1973).

Cloning and expression of human lipocortin, a phospholipase A₂ inhibitor with potential anti-inflammatory activity

Barbara P. Wallner, Robert J. Mattaliano, Catherine Hession, Richard L. Cate, Richard Tizard, Lesley K. Sinclair, Carl Foeller, E. Pingchang Chow, Jeffrey L. Browning, K. L. Ramachandran & R. Blake Pepinsky

Biogen Research Corp., 14 Cambridge Center, Cambridge, Massachusetts 02142, USA

The anti-inflammatory action of glucocorticoids has been attributed to the induction of a group of phospholipase A₂ inhibitory proteins, collectively called lipocortin. These proteins are thought to control the biosynthesis of the potent mediators of inflammation, prostaglandins and leukotrienes, by inhibiting the release of their common precursor, arachidonic acid, a process that requires phospholipase A₂ hydrolysis of phospholipids¹⁻¹¹. Lipocortin-like proteins have been isolated from various cell types, including monocytes³, neutrophils⁴ and renal medullary⁹ cell preparations. The predominant active form is a protein with an apparent relative molecular mass (M_r) of 40,000 (40K). These partially purified preparations of lipocortin mimic the effect of steroids, and mediate anti-inflammatory activity in various *in vivo* model systems^{3,5,11}. Using amino-acid sequence information obtained from purified rat lipocortin, we have now cloned human lipocortin complementary

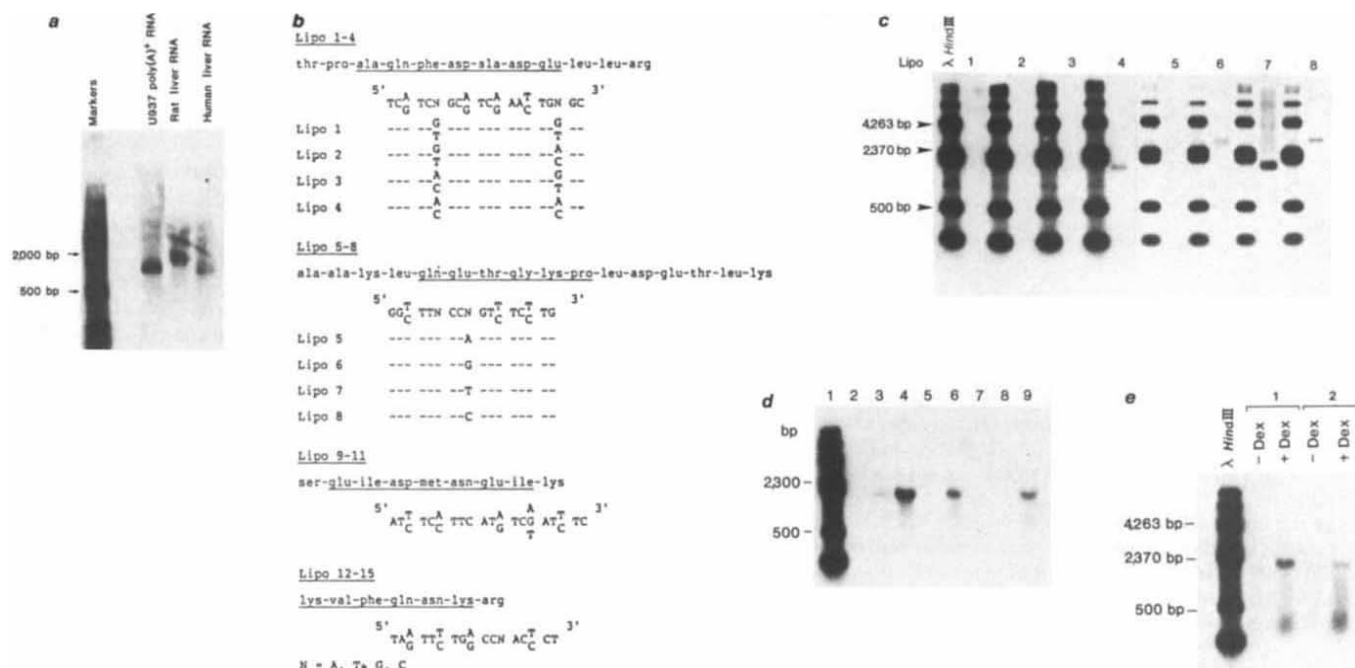


Fig. 1 *a*, Northern blot analysis of U937, rat liver and human liver RNA to determine homology between the oligonucleotide sequence (derived from rat lipocortin) and human RNA. Oligonucleotide pool Lipo 1-4 was used as hybridization probe. *b*, The four pools of oligomers synthesized. *c*, Northern blot analysis on U937 poly(A) RNA to reduce the degeneracy of the hybridization probes Lipo 1-4 and Lipo 5-8. Numbers on the top line indicate the number of the oligonucleotide subpools used as hybridization probes. λ DNA digested with HindIII was used as markers. *d*, Northern blot analysis of total RNA extracted from rat organs: lane 1, λ HindIII markers; 2, liver; 3, heart; 4, lung; 5, brain; 6, thymus; 7, muscle; 8, kidney; 9, spleen. *e*, Northern blot analysis of total RNA from rat peritoneal cells: lanes 1, uninduced (−Dex) and dexamethasone-induced (+Dex); 2, uninduced and dexamethasone-induced after treatment with mineral oil.

Methods. Fragments of tryptic digests of purified rat lipocortin resolved by reverse-phase HPLC on a C₁₈ column were subjected to amino-terminal protein sequence analysis on an applied Biosystems 470A sequencer. Phenylthiohydantoin-amino acids from each cycle of the Edman degradation were analysed by reverse-phase HPLC on a Cyano column. From the sequences of the peptides shown in *b*, four pools of oligomers were chemically synthesized¹⁶ using a 380A Applied Biosystem DNA synthesizer. U937 cells were induced with 10^{−5} M dexamethasone and 10^{−7} M phorbol ester (4β-phorbol 12β-myristate 13α-acetate; Sigma), pelleted, resuspended in guanidinium thiocyanate buffer, lysed by vortexing, and the RNA was extracted as described previously¹⁷. Total RNA was chromatographed on an oligo(dT)-cellulose column to enrich for poly(A) RNA. Total rat liver and total human liver RNA were isolated as described elsewhere¹⁷. Northern blot analysis: 2 μg of poly(A) RNA from U937 cells and 20 μg of total rat liver and human liver RNA were electrophoresed in a formaldehyde-1% agarose gel by a modified procedure of Maniatis *et al.*¹⁸. The gel consisted of 1×MOPS buffer (40 mM MOPS pH 7.0, 10 mM sodium acetate, 1 mM EDTA pH 7.0), 0.74 M formaldehyde and was electrophoresed in 1×MOPS buffer at 100 V for 2 h. The loading buffer was 1.8 M formaldehyde, 1×MOPS buffer, 50% formamide, 0.01% bromophenol blue, 0.01% xylene cyanole FF, 3% Ficoll. RNA samples were denatured at 65 °C for 15 min before loading. After electrophoresis the gel was rinsed five times in distilled H₂O for 5 min each, soaked in 20×SSC (3 M sodium chloride; 0.3 M sodium citrate, pH 7.0) for 1 h and transferred to GeneScreen membrane in 2×SSC (0.3 M sodium chloride; 0.03 M sodium citrate, pH 7.0) for 6 h. After transfer, the RNA was linked to the membrane by ultraviolet treatment as described previously¹⁹ for 3 min. Prehybridization was performed in 5×SSC (0.75 M sodium chloride, 0.075 M sodium citrate pH 7.0), 50 mM sodium phosphate pH 7.0, 10×Denhardt's (0.2% polyvinylpyrrolidone (*M_n* 40,000), 0.2% Ficoll (*M_n* 400,000), 0.2% bovine serum albumin), 100 μg ml^{−1} transfer RNA, 7% SDS, at 45 °C for 1 h. Hybridization was in 10% dextran sulphate, 5×SSC, 10×Denhardt's, 20 mM sodium phosphate pH 7.0, 100 μg ml^{−1} tRNA, 7% SDS, oligonucleotide pool Lipo 1-4, ³²P-labelled by T4 polynucleotide kinase (specific activity 1×10⁷ c.p.m. pmol^{−1}) at a concentration of 1×10⁵ c.p.m. per ml of hybridization buffer, at 45 °C for 18 h. Washes were performed at 45 °C in 1×SSC, 0.5% SDS. *c*, Poly(A)⁺ RNA extracted from U937 cells was electrophoresed on a 1% formaldehyde agarose gel and transferred to GeneScreen membrane, which was cut into eight individual strips, each containing one lane of RNA and one lane marker. Each strip was hybridized individually in 2 ml of hybridization buffer as described above with the ³²P-labelled oligonucleotide probe as indicated (specific activity 1×10⁷ c.p.m. pmol^{−1}) at a concentration of 1×10⁵ c.p.m. ml^{−1} of buffer at 45 °C for Lipo 1-4 and 50 °C for Lipo 5-8. All filters were washed with 1×SSC, 0.5% SDS at the hybridization temperature. *d*, *e*, Rat resident peritoneal cells were isolated by washing the peritoneal cavity with 25 ml of 0.34 M sucrose and cells pelleted by centrifugation for 5 min at 2,500g. In experiments with dexamethasone, the rats were pretreated i.p. with 1.25 mg kg^{−1} dexamethasone phosphate in normal saline and cells were collected after 2 h. For oil-elicited peritoneal cells, rats were injected i.p. with 8 ml of Fisher lightweight paraffin oil and the cells isolated after 3 days. Extraction of total RNA from rat organs and peritoneal cells, and Northern blot analysis, were performed as described above, except that a restriction fragment of λL4-211 (³²P-labelled by nick-translation) was used as a probe.

DNA and expressed the gene in *Escherichia coli*. Our studies confirm that lipocortin is a potent inhibitor of phospholipase A₂ activity.

A 37K phospholipase A₂ inhibitory protein was purified from rat peritoneal exudates and partially sequenced¹². Amino-acid sequence information from four of the tryptic peptides was used to chemically synthesize pools of oligonucleotide probes (Fig. 1). Northern analysis of poly(A) RNA from the human histiocytic lymphoma cell line U937, human liver and rat liver showed a distinct transcript of 1,400–1,600 nucleotides in each messenger RNA preparation hybridizing with oligonucleotide probes Lipo 1-4 (Fig. 1*a*) and Lipo 5-8 (data not shown). These data suggested that rat and human lipocortin sequences are sufficiently homologous for the oligonucleotides derived from the rat sequence to be used to screen for the human gene directly. To reduce the degeneracy of the oligonucleotide probe pools, subpools were hybridized individually to U937 mRNA in a Northern blot analysis (Fig. 1*c*). This approach allowed selection of a single subpool from each set of probes and thus decreased

the degeneracy for Lipo 1-4 from 256 to 64 and that for Lipo 5-8 from 128 to 32.

A U937 cDNA library was constructed in the vector λgt10 and screened with ³²P-labelled oligonucleotide probe subpool Lipo 4. Isolated positive recombinant phages were characterized further by hybridization to all four oligonucleotide probe pools. DNA sequence determination of the largest insert (λL4-211) has an open region reading frame of 1,038 base pairs (bp) and a 3'-untranslated region of 283 bp which includes a poly(A) addition site (Fig. 2). All four rat lipocortin polypeptides, from which the oligonucleotide probes were derived, were located within the amino-acid sequence deduced from the human cDNA sequence (Fig. 2). The translational initiation codon was assigned to the first in-frame ATG. Since the amino-terminus of the rat protein was blocked, we could not deduce the translational start site. The length and nucleotide sequence of the 5'-untranslated sequence of lipocortin mRNA were determined by experiments outlined in Fig. 3. Analysis of the 5' fragment generated by digestion of a lipocortin mRNA/oligonucleotide

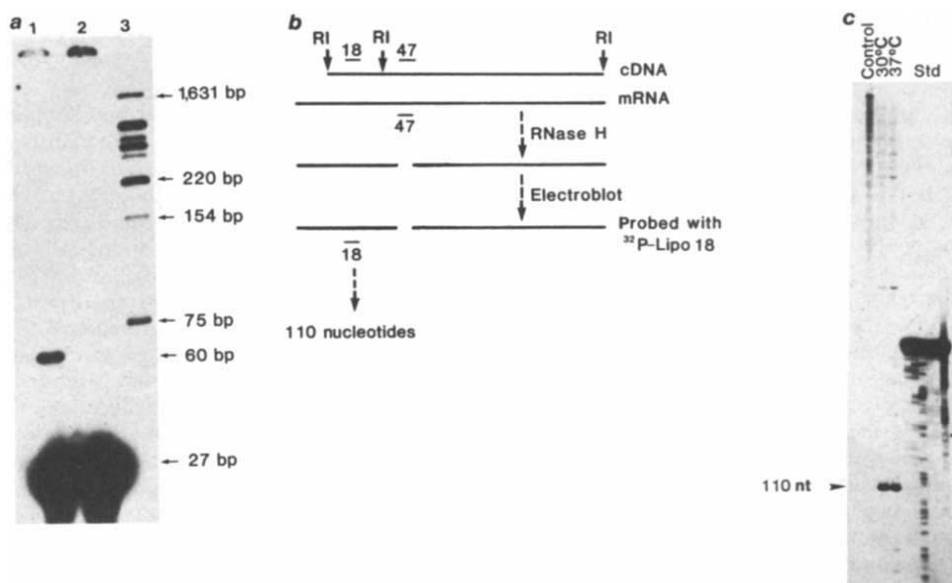
Fig. 2 Nucleotide sequence and predicted amino-acid sequence of human lipocortin. The sequence printed in lower-case letters is the mRNA sequence determined by primer extension, as described in Fig. 3 legend. The nucleotide sequence printed in capital letters represents the cloned cDNA sequence. Nucleotides are numbered from the presumed initiator ATG. The amino-acid sequence of the rat lipocortin peptides determined by amino-acid sequencing is printed for comparison above the deduced amino-acid sequence for human lipocortin. Amino acids that differ between the rat and human sequence are underlined in the rat sequence.

Methods. For the construction of the U937 cDNA library, poly(A) RNA was isolated from U937 cells induced with dexamethasone and phorbol ester (4β -phorbol 12 β -myristate 13 α -acetate) as described for Fig. 1. First- and second-strand cDNA was synthesized as described previously¹⁸. Double-stranded cDNA of ≥ 500 bp was inserted into λ gt10 arms as described previously²⁰. The resulting cDNA library of 1×10^6 independent recombinants was screened with 32 P-5'-labelled oligonucleotide probe subpool Lipo 4 (specific activity 1×10^8 c.p.m. pmol⁻¹) at a concentration of 1×10^5 c.p.m. per ml hybridization buffer, as described previously²¹. Recombinant λ L4-211, which cross-hybridized to all four oligonucleotide probe pools, was isolated and subcloned into pUC13 (pLip 4/10). The nucleotide sequence was determined as described previously²². For comparison of 5' cDNA sequences, an *Eco*RI/*Bgl*II restriction fragment of pLip 4/10 was 32 P-labelled by nick-translation and used to rescreen the U937 cDNA library. λ DNA minipreparations of 36 isolated positives were digested with *Msp*I, and the DNA cleaved by Maxam and Gilbert chemistry and electroblotted onto GeneScreen after separation on a urea/polyacrylamide gel as described previously¹⁹. The nucleotide sequence determined after hybridization of the filter to a 32 P-5'-end-labelled oligonucleotide probe homologous to the sequence 44–78 of the coding sequence, showed that none of the 36 clones had sequences additional to that contained in λ L4-211. Details and applications of this technique are described elsewhere (R.T. and H. Nick, manuscript in preparation).

Human	-74	agtggtgaatcttcagagaagaaTTCTCTTTAGTTCTTTCGCAAGAGGTAGAGATAAAGACACTTTTCAAAAATGGCAATGGTATCAGAAATCTCTCAA	26
Human	27	GlnAlaTyrPheIleGluAsnGluGluGlnGlnTyrValGlnThrValLysSerSerLysGlyGlyProGlySerAlaValSerProTyrProThrPhe	126
Human	127	GCAGGCTCGGTTATTGAAATGAAGAGCAGGAATATGTTCAAACCTGTGAAGTCATCCAAAGGTGGTCCCGATCAGCGGTGAGCCCTATCTCACTTC	
Human	127	AsnProSerSerAspValAlaAlaLeuHisLysAlaIleMetValLysGlyValAspGluAlaThrIleIleAspIleLeuThrLysArgAsnAsnAla	226
Rat	127	AATCCATCTCGGATGCGTGCCTTGATAAGCCATAATGGTTAAAGGTGGATGAAGCAACCATCATTTGACATTTCTAACTAAGCGAAACATTCAC	
Human	227	GlnArgGlnGlnIleLysAlaIleTyrLeuGlnGlnTyrGlyLysProLeuAspGluThrLeuLys	326
Human	227	AGCGTCAACAGATCAAGCAGCATATCTCAGGAACAGGAAAGCCCTGGATGAACACTTAAGAAAGCCCTACAGGTACCTTGAGGAGGTGTTT	
Rat		Lipo5-8	
Rat		LeuLysThrProAlaGlnPheAspAlaAspGluLeuLeuArg	
Human	327	AlaLeuLeuLysThrProAlaGlnPheAspAlaAspGluLeuArgAlaIleMetLysGlyLeuGlyThrAspGluAspThrLeuIleGluIleLeuAla	426
Human	327	AGCTCTGCTAAAACCTCCAGCGCAATTTGATGCTGTAAGCTCTGCTGCTGATGAAGGCTTGGAACTGATGAAGTACTCTAATTGAGATTTTGCA	
Human	427	Lipo1-4	
Human	427	SerArgThrAsnLysGluIleArgAspIleAsnArgValTyrArgGluGluLeuLysArgAspLeuAlaLysArgIleThrSerAspThrSerGlyAsp	526
Human	427	TCAGAACTAACAAAGAAATCAGAGACATTAACAGGGTCTCAGAGAGGAACCTGAAGAGCACTGGCCAAAGACATAACCTCAGACACATCTGGAGATT	
Human	527	PheArgAsnAlaLeuLeuSerLeuAlaLysGlyAspArgSerGluAspPheGlyValAsnGluAspLeuAlaAspSerAspAlaArgAlaLeuTyrGluAla	626
Human	527	TTCCGAACGCTTGTCTTCTCTTCTGCTAAGGGTGACCGACTGAGGACTTTGGTGGTAAATGAAGACTTGGCTGATTCAGATGCCAGGGCTTGTATGAAG	
Rat		LysValPheGlnAsnTyrArg	
Human	627	GlyGluArgArgLysGlyThrAspValAsnValPheAsnThrIleLeuThrThrArgSerTyrProGlnLeuArgArgValPheGlnLysTyrThrLys	726
Human	627	AGGAGAAAGGAGAAAGGGGAGCAGCTAAACGCTGTTCAATACCATCTTACCACCAAGACTTCCCAACTTCGACAGTGTTCAGAAATACACCAAG	
Human	727	Lipo12-15	
Human	727	TyrSerLysHisAspMetAsnLysValLeuAspLeuGluLeuLysGlyAspIleGluLysCysLeuThrAlaIleValLysCysAlaThrSerLysPro	826
Human	727	TACAGTAAGCATGACATGAACAAAGTCTGACCTGGAGTGAAGGTGACATTGAGAAATGCCTCAGCATATCGTGAAGTGCACCAAGCAACACAG	
Rat		SerGluIleAspMetAsn	
Human	827	AlaPhePheAlaGluLysLeuHisGlnAlaMetLysGlyValGlyThrArgHisLysAlaLeuIleArgIleMetValSerArgSerGluIleAspMetAsn	926
Human	827	CTTCTTTTGCAGAGAGCTTATCAAGCCATGAAGGCTTGGAACTCGCCATAGGCATTGATCAGGATTATGGTTCCCGTCTCGAAATGACATGAA	
Rat		Lipo9-11	
Rat		GluIleLys	
Human	927	AspIleLysAlaPheTyrGlnLysMetTyrGlyIleSerLeuCysGlnAlaIleLeuAspGluThrLysGlyAspTyrGluLysIleLeuValAlaLeu	1026
Human	927	TGATATCAAGCATCTATCAGAGATGATGTTATCTCCCTTTGCCAAGCCATCTCGGATGAACCAAGAGAGATTAGAGAAATCTGTGGCTCTT	
Human	1027	CysGlyGlyAsn	1126
Human	1027	TGTGAGGAACATAACATTCCTTGATGCTCAAGCTATGATCAGAACTTAATATATATTTTCATCCTATAAGCTTAATAGGAAAGTTCTTC	
Human	1127	AACAGGATTACAGTGTAGCTACCTACATGCTGAAAAATATAGCCTTTAAATCATTTTATATTATAACTCTGATAATAGAGATAAGTCCATTTTATAA	1226
Human	1227	AATGTTTTCCCAACCAATAAACCCATACAAAGTGTCTAGTAAACATCATGAGAAGATGTCTATGTAGCTGAAATATAATGACGTACAAGACA	1326
Human	1327	polyA signal	
Human	1327	AAAAAAAAA	

Fig. 3 Determination of the 5' lipocortin mRNA sequence. *a*, Primer extension analysis of human placental poly(A) RNA (lane 1) and human liver poly(A) RNA (lane 2). Lane 3, *Hinf*I-digested pBR322. *b*, Schematic outline of the procedure to determine the size of the lipocortin mRNA using RNase H. *c*, Gel analysis of the 5' mRNA fragments obtained by digestion of the mRNA-Lipo 47 hybrid with RNase H. The control lane contains RNase H digestion of placental poly(A) RNA without oligonucleotide Lipo 47 present. The second and third lanes contain digests of the poly(A) RNA-Lipo 47 hybrid, incubated with RNase H at 30 or 37 °C. The size marker is a DNA fragment of known sequence cleaved by Maxam and Gilbert sequencing chemistry.

Methods. Primer extension: 4 μ g of placental poly(A) RNA was hybridized to 0.2 pmol of kinased oligonucleotide Lipo 18 (homologous to base pairs -14 to -41 of the sequence shown in Fig. 2) for 90 min at 65 °C in 0.04 M PIPES (pH 6.4), 0.4 M NaCl and 1 mM EDTA, ethanol-precipitated and resuspended in 50 mM Tris (pH 8.3), 20 mM MgCl₂, 6 mM β -mercaptoethanol, 2 mM deoxynucleotide phosphates and 1 U ml⁻¹ RNasin. Avian myeloblastosis virus reverse transcriptase (19 U) was added and the reaction incubated at 41 °C for 60 min, ethanol-precipitated and resuspended in 10 μ l of a formamide loading buffer. The extension products were denatured and separated on a 6% acrylamide gel containing 7 M urea. The 60-bp band was eluted and its nucleotide sequence determined²². RNase H digestion: Human placental poly(A) RNA was prepared as described previously¹⁷. Poly(A) placental RNA (5 μ g) was hybridized to 1.2 pmol oligonucleotide Lipo 47 (homologous to base pairs 27–37 of lipocortin coding sequence) in 40 mM Tris-HCl pH 7.9, 4 mM MgCl₂, 1 mM dithiothreitol at 65 °C for 10 min, then at 30 or 37 °C for 45 min. Digestions were carried out with 10 units of RNase H for 1.5 h at 30 or 37 °C. RNA samples were denatured at 65 °C and sized on a 7 M urea, 6% polyacrylamide gel, transferred electrophoretically to a Biotodyne A nylon membrane and hybridized to oligonucleotide probe Lipo 18 (homologous to nucleotides -14 to -41 of the sequence presented in Fig. 2), which had been 32 P-labelled by T4 polynucleotide kinase and [γ - 32 P]ATP. The prominent band co-migrates with a marker DNA sequence band of 110 nucleotides (nt). The oligonucleotide probe Lipo 47 is homologous to base pairs 78–88 of the cloned sequence. Assuming that RNase H clips the hybrid at nucleotide 83, the size of the mRNA fragment 5' to this sequence is 27 nucleotides, which agrees well with the 23 nucleotides obtained by sequencing the primer-extended fragment.



hybrid with RNase H demonstrated that we had cloned all but 27 nucleotides of the lipocortin mRNA (Fig. 3b, c). Primer extension of the message and sequencing of the extended cDNA also showed an additional 23–27 nucleotides not found in the cloned sequence (Fig. 3a). This sequence did not contain a termination codon or another ATG upstream of the ATG described above.

Human lipocortin is a strongly polar molecule; ~30% of its amino acids are charged residues, which are distributed throughout the molecule. The deduced amino-terminal sequence of the recombinant human lipocortin does not contain a leader peptide, but there are short stretches of internal hydrophobic residues (Fig. 2). Although lipocortin is obtained as an extracellular protein from rat peritoneal lavages, the mechanism that promotes the release of this molecule from macrophages is unknown. Lipocortin contains a single potential glycosylation site, Asn-Pro-Ser, located in the nucleotide sequence at base pairs 121–129. The protein also contains the sequences Glu-Glu-Gln-Glu-Tyr located at position 49–63, and Arg-Arg-Lys-Gly-Thr at base pairs 634–648, which are, respectively, potential tyrosine and threonine phosphorylation sites^{13,14}. This is consistent with published observations that lipocortin can be phosphorylated by tyrosine and serine threonine kinases^{13,14}. Although recombinant human lipocortin can be phosphorylated *in vitro* by both tyrosine and serine kinases (L. Cantley, personal communication), it is not known whether the sites of phosphorylation correspond to those above.

To confirm that the isolated lipocortin cDNA encodes a phospholipase A₂ inhibitory protein, the full-length coding sequence was expressed in *E. coli*. Plasmid pLiptrc155A (Fig. 4c) contains the coding sequence of the lipocortin gene under the regulation of the *trc* promoter¹⁵ and was designed so that translation would initiate at the first ATG of human lipocortin mRNA. When crude cell lysates from *E. coli* strains transformed with pLiptrc155A were analysed by SDS-polyacrylamide gel electrophoresis, a single protein of apparent *M*_r 37K, specific to cells containing the lipocortin gene, was detected (Fig. 4a). In *E. coli* JA221 it represents about 4% of the total protein. Western blot analysis of a parallel set of samples showed that only cells harbouring pLiptrc155A contain an immunoreactive protein (Fig. 4b). Amino-terminal micro-sequencing and amino-acid analysis (not shown) further confirm that the immunoreactive 37K protein is encoded by the cloned lipocortin gene.

Recombinant lipocortin is active as a phospholipase A₂ inhibitor, and this activity can be measured directly in crude lysates of *E. coli* expressing lipocortin (Fig. 4a legend). When these cultures are lysed and the soluble fractions assayed, phospholipase A₂ inhibitory activity is detected specifically in those preparations where lipocortin is expressed. The level of inhibitory activity correlates with expression levels as determined by SDS-polyacrylamide gel electrophoresis. This activity is trypsin-sensitive and is destroyed by boiling (data not shown). The purification and partial sequence of rat lipocortin have been reported elsewhere¹². A comparison of the properties and tryptic map of the human recombinant lipocortin with the purified rat protein shows that these two proteins are very similar. Their amino-acid sequences are highly conserved, most of the differences occurring at the amino- and carboxy-termini. Antibodies raised against purified rat lipocortin¹² cross-react with the human recombinant protein. We estimate that purified recombinant lipocortin has a specific activity of ~6,000 U per mg, which is comparable to activities detected with natural rat protein¹².

Using Northern analysis, we determined lipocortin mRNA levels in rat peritoneal macrophages and several rat organs. High concentrations are detected in lung, thymus and spleen (Fig. 1d), while liver, heart, brain, muscle and kidney show small but detectable amounts of lipocortin mRNA. A similar distribution of lipocortin is seen by Western analysis (data not shown). Also using Northern analysis, we have determined that

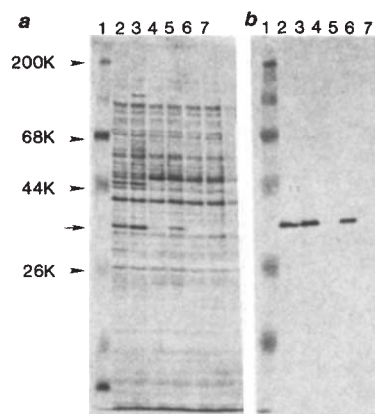


Fig. 4 Expression of lipocortin in *E. coli*. **a**, SDS-polyacrylamide gel analysis of total cell lysates visualized by Coomassie staining. Lane 1, BRL prestained high-relative molecular mass markers; 2, pLiptrc155A in JA221 uninduced (47% phospholipase A₂ inhibitory activity in crude *E. coli* lysate); 3, induced with 1 mM isopropyl thiogalactoside (IPTG, 57%); 4, 5, pLiptrc155A in W3110I^Q uninduced (2%) and induced (24%); 6, 7, W3110I^Q control (no lipocortin gene), uninduced (0%) and induced (2%). **b**, Western blot analysis of cell lysates, separated on an SDS-polyacrylamide gel as in **a**. **c**, Lipocortin expression vector.

Methods. **a, b**, Cultures of strains indicated were grown in L-broth to A₅₅₀ of 1.0 and induced with 1 mM IPTG for 1 h at 37 °C. Cell pellets were dissolved in electrophoresis sample buffer containing SDS and β-mercaptoethanol and analysed on a 12.5% SDS-polyacrylamide gel. Protein bands were visualized either directly by Coomassie staining or by Western blot analysis with an antiserum raised against SDS-disrupted rat lipocortin. For the phospholipase A₂ inhibitory assay, cultures of strains indicated were pelleted by centrifugation and resuspended at 40 A units per ml in 25 mM Tris-HCl pH 7.7, 5 mM EDTA. Cells were lysed with a French pressure cell at 12,000 p.s.i. and particulate debris removed by centrifugation (10 min, 12,000 r.p.m. in a Sorvall SS34 rotor). Aliquots (15 μl) of each lysate were assayed for phospholipase A₂ inhibitory activity¹⁰. **c**, For the construction of plasmid pLiptrc155A, plasmid pLip4/10 was digested with *Eco*RI and partially with *Hind*III. The gel-purified fragment was inserted into the *Nco*I/*Hind*III-digested expression vector pKK233-2 (ref. 15) together with a synthetic *Nco*I/*Eco*RI linker. The *Nco*I/*Eco*RI linker

```
5' CATGGCAATGGTATCAG
3' CGTTACCATAGTCTTAA
```

replaces the missing coding region for the first 5 amino acids, starting with the presumed initiating ATG for lipocortin.

human placenta contains ~100-fold higher levels of lipocortin mRNA than human liver (data not shown).

Induction of lipocortin by glucocorticoids has been reported in various tissues and resident peritoneal exudate cells. We find that when rats are treated with dexamethasone *in vivo*, lipocortin mRNA is induced. This result is particularly apparent in the analysis of resident rat peritoneal cells (Fig. 1e). Two hours after an intraperitoneal (i.p.) injection of dexamethasone, the concentration of mRNA in peritoneal cells is sixfold higher than that detected in untreated rats. When mineral oil-elicited peritoneal exudate cells were analysed in a similar manner (Fig. 1e), dexamethasone caused only a slight increase in transcription over untreated controls. In the same set of experiments, dexamethasone had no effect on the intracellular level of the 37K protein detected by Western analysis (not shown). We cannot explain this phenomenon, although similar increases in transcription of lipocortin mRNA were observed on dexamethasone treatment of lymphoma U937 cells or human monocytes (not shown). These results indicate that the activity of the gene can be modulated by glucocorticoids, and support previous observations which suggest that steroids induce lipocortin production.

Southern blot analysis of human genomic DNA using two different oligonucleotides homologous to the 5' sequence of cloned DNA indicates that lipocortin is a unique gene (not shown). Both oligomers detected only one restriction fragment per digest. Southern analysis using nick-translated probes from the cDNA is also consistent with this interpretation. In addition,

Northern analyses of mRNA preparations from various tissues showed only one major lipocortin mRNA species. These data are consistent with the hypothesis that other phospholipase A₂ inhibitory proteins in this family are either proteolytic cleavage products or aggregates of the 37K protein. The availability of the purified recombinant lipocortin will aid in characterizing physical properties of the protein, and will increase our understanding of its biological role in inflammation.

We thank Richard Fisher, Harry Meade, Harry Nick, Gary Gray, Tim Kovacic, Donna Costopoulos, Liora Lieberman, Rande Leibowitz, Patti Devlin and Tanya Falbel for their contribution to this work; Karen Nastasia, Rosemary Morrissey and Susan Pagan for their assistance in DNA sequencing; and Joanne Jablonski and Leslie Wachter for their preparation of the oligonucleotides. We appreciate the continued support and helpful discussions of Richard A. Flavell and John E. Smart. This work was supported by Biogen Research Corp. and Yamanouchi Pharmaceutical Company.

Received 23 September 1985; accepted 10 January 1986.

- Blackwell, G. J. *et al. Nature* **287**, 147-149 (1980).
- Carnuccio, R., DiRosa, M. & Persico, P. *Br. J. Pharmac.* **68**, 14-16 (1980).
- Flower, R. J. & Blackwell, G. J. *Nature* **278**, 456-459 (1979).
- Hirata, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 2533-2536 (1980).
- Cloix, J. F. *et al. Br. J. Pharmac.* **79**, 313-321 (1983).
- Hong, S.-C. & Levine, L. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1720-1734 (1976).
- Blackwell, G. J. *et al. Br. J. Pharmac.* **62**, 79-89 (1978).
- Carnuccio, R. *et al. Br. J. Pharmac.* **74**, 322-324 (1981).
- DiRosa, M. & Persico, P. *Br. J. Pharmac.* **66**, 161-163 (1979).
- Rothhut, B. *et al. Biochem. biophys. Res. Commun.* **117**, 878-884 (1983).
- Flower, R. J., Wood, J. N. & Parente, L. *Advances in Inflammation Research* Vol. 7 (eds Otterness, I., Capetola, R. & Wong, S.) 61-70 (Raven, New York, 1984).
- Pepinsky, R. B. *et al. J. biol. Chem.* (in the press).
- Hirata, F. *J. biol. Chem.* **256**, 7730-7733 (1981).
- Hirata, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 4717-4721 (1984).
- Amann, E. & Brosius, J. *Gene* **40**, 183-190 (1985).
- Matteucci, M. D. & Carutiers, M. H. *J. Am. chem. Soc.* **103**, 3185-3191 (1981).
- Chirgwin, J. M. *et al. Biochemistry* **18**, 5294-5299 (1979).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning*, 229-246 (Cold Spring Harbor Laboratory, New York, 1982).
- Church, G. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1991-1995 (1984).
- Huynh, T., Young, R. A. & Davis, R. W. *DNA Cloning Techniques, A Practical Approach* (ed. Glover, D. M.) 49-78 (IRL Press, Oxford, 1985).
- Woo, S. L. *Meth. Enzym.* **68**, 389-395 (1979).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 497-560 (1980).

Removal of the *Alu* structural domain from signal recognition particle leaves its protein translocation activity intact

Vivian Siegel & Peter Walter

Department of Biochemistry and Biophysics,
University of California, San Francisco, California 94143, USA

Alu-like elements comprise the most abundant family of interspersed repetitive sequences in primates and rodents¹, and contain many features of processed genes², suggesting that they were initially derived by reverse transcription of processed RNA transcripts. Transcripts containing *Alu* family members are represented in heterologous nuclear RNAs, cytoplasmic messenger RNAs and small RNAs^{1,3}, although nothing is known about their function. Evolutionary studies^{4,5} strongly suggest that the parent RNA for the *Alu*-like elements is the highly conserved 7SL RNA, which is an essential component of signal recognition particle (SRP), a small cytoplasmic ribonucleoprotein whose function is the targeting of nascent secretory and membrane proteins to the rough endoplasmic reticulum (for a review see ref. 6). 7SL RNA is composed of both unique and *Alu*-like sequences⁷. SRP is rod-shaped⁸ and, in addition to its RNA, contains four proteins (two monomers composed of a polypeptide of relative molecular mass (M_r) 19,000 (19K) and one of 54K, and two heterodimers, one composed of a 9K and a 14K polypeptide, and the other

composed of a 68K and a 72K polypeptide, respectively)^{9,10}. The RNA moiety is required for SRP activity, as well as for structural integrity of the particle⁴. To investigate whether the *Alu*-like segments of 7SL RNA have a specific role in SRP activity, we have now purified and analysed a SRP subparticle that is created upon extensive digestion with micrococcal nuclease and entirely lacks the *Alu*-like sequences. We find that it contains the 72/68K, 54K and 19K proteins tightly bound, but lacks the 9/14K protein. *In vitro* activity assays demonstrated that the subparticle could still promote secretory protein translocation across the microsomal membrane, but could no longer trigger an arrest of pre-secretory protein synthesis. Re-addition of the 9/14K protein did not restore the elongation arrest. We conclude that the region of SRP comprised of the *Alu*-like RNA and the 9/14K protein exists in a distinct structural domain which is not required for the protein translocation promoted by SRP but apparently confers elongation-arresting activity on the particle.

Approximately 100 nucleotides at the 5' end and 45 nucleotides at the 3' end of 7SL RNA are homologous to the human *Alu* right monomer sequence⁷, and, in proposed secondary structures, the 5' and 3' segments base-pair with each other such as to form one end in the folded RNA (see below). It had previously been shown that the RNA in SRP can be cleaved quite selectively with micrococcal nuclease such that the central unique 'S segment', which is 155 nucleotides long and shows no homology to *Alu* DNA, remains largely intact and is separated from the *Alu*-like segments¹¹. SRP was digested with a 10-fold molar excess of micrococcal nuclease to yield five major RNA fragments (Fig. 1a). We used DEAE-Sepharose chromatography and sucrose gradient sedimentation to purify to homogeneity a subparticle containing the major 150-nucleotide fragment (see Fig. 1). The sucrose gradient fraction containing this fragment was pure with respect to RNA content both by gel analysis of end-labelled RNA and by primer extension to map the 5' end of the molecule (see Fig. 1 legend); we found that the RNA present in this fraction consisted entirely of the unique S sequences of 7SL RNA. In addition, the fraction contained the 68/72K, 54K and 19K proteins, but had lost the 9/14K protein. We shall refer to the subparticle as SRP(S).

SRP has been shown to have two separable activities which can be assayed in a wheat-germ cell-free translation system^{6,12}. First, it promotes the translocation of secretory proteins across the membrane of SRP-depleted microsomes ('translocation-promoting' activity), which can be evidenced by signal peptidase cleavage; without SRP, such translocation and processing do not occur. Second, SRP selectively inhibits the synthesis of pre-secretory proteins at the level of elongation ('elongation-arresting' activity).

We have measured the translocation-promoting and elongation-arresting activities of SRP(S). Figure 2 demonstrates that SRP(S) promotes the translocation of preprolactin across the microsomal membrane (Fig. 2a), although ~60% less efficiently than the same concentration of undigested SRP. We do not know whether the decreased activity is a result of an inactivation that occurred during the additional fractionation, or whether it is an intrinsic property of the subparticle. In marked contrast to its ability to promote secretory protein translocation, however, SRP(S) was unable to trigger a detectable arrest of preprolactin synthesis (Fig. 2b). Figure 3 confirms these results.

Translocation across the membrane of the endoplasmic reticulum seems to occur in a co-translational manner, and the elongation-arresting activity of SRP has been shown *in vitro* to maintain the nascent chain in a translocation-competent state; that is, when SRP is present in a synchronized translation, it will block elongation, so that microsomal membranes can be added at a later time and translocation will still occur¹³ (Fig. 3a, solid squares). However, when SRP(S) is present in a synchronized translation (Fig. 3a), translocation is strictly dependent on the time at which membranes are added. Thus, SRP(S) lacks the elongation-arresting activity required to maintain the

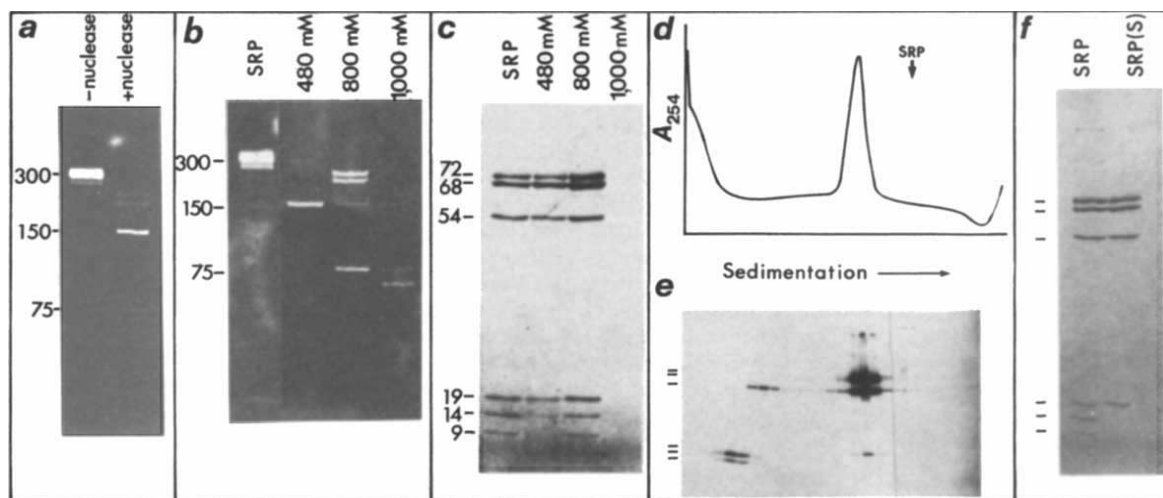


Fig. 1 Purification of SRP(S). Gradient-purified SRP (200 μ g in 1 ml, purified from canine pancreas as described elsewhere¹⁷) was digested with 4,000 U *Staphylococcus aureus* nuclease (Boehringer) in a solution containing 50 mM triethanolamine/HOAc pH 7.5, 250 mM KOAc, 2 mM Mg(OAc)₂, 1 mM CaCl₂, 1 mM dithiothreitol (DTT) and 0.01% Nikkol (octa-ethyleneglycol-mono-*n*-dodecyl ether, Nikko Chemicals, Japan). The digestion was allowed to proceed for 1 h at 30 °C and was then stopped by the addition of EGTA to 3.8 mM. **a**, An ethidium bromide-stained 7 M urea/6% polyacrylamide gel of RNA prepared from SRP incubated without and with micrococcal nuclease. Note that the major product of the digestion is the 150-nucleotide S fragment. The rest of the purification was performed at 0–4 °C. The nuclease-digested material was loaded directly onto a 250- μ l DEAE-Sepharose CL-6B column. The column was washed with 10 column vols 20 mM HEPES/KOH pH 7.5, 2 mM Mg(OAc)₂, 2 mM EGTA and 0.01% Nikkol (buffer A), containing 250 mM KOAc. It was then sequentially eluted with 530- μ l aliquots of buffer A containing the following concentrations of KOAc: 480 mM, 800 mM and 1,000 mM. One 80- μ l fraction and six 125- μ l fractions were collected at each step. Fractions 2, 3 and 4 were pooled and the protein and RNA components displayed by polyacrylamide gel electrophoresis (PAGE) followed by Coomassie blue or ethidium bromide staining, respectively. **b**, **c**, The RNA and protein compositions of the eluted fractions at the indicated KOAc concentrations. Note that the S fragment elutes at a lower ionic strength than any of the other fragments, probably because the particle has a higher protein to RNA ratio. All four proteins are present in both the 480 mM and the 800 mM salt steps, although the 9/14K protein is present in lower amounts in the 480 mM step than in the starting SRP, and no protein is detectable in the 1,000 mM step (see below). The peak from the 480 mM KOAc elution was loaded onto a 5–20% sucrose gradient (12 ml) in 20 mM HEPES/KOH pH 7.5, 500 mM KOAc, 5 mM Mg(OAc)₂, 1 mM EGTA, 1 mM DTT and 0.01% Nikkol. The gradients were centrifuged in a Beckman SW40 swinging bucket rotor for 24 h at 40,000 r.p.m. Gradients were fractionated with an Isco gradient fractionator and 0.5-ml fractions collected. **d**, The absorbance profile of the gradient monitored at 254 nm. The left side is the top of the gradient. **e**, The polyelectrolyte profile of an analytical gradient spun and fractionated in identical conditions to those described above, but using as starting material in the preparation SRP that had been radioactively labelled with ¹²⁵I-Bolton-Hunter reagent¹³. The polypeptide components of each of the fractions were separated by SDS-PAGE on 10–15% gradient gels and visualized by autoradiography. A preparative gradient run in parallel was monitored by SDS-PAGE and Coomassie blue staining and shows an identical profile. Note that the 9/14K protein stays at the top of the gradient, whereas the 68/72K, 54K and 19K proteins co-sediment with the RNA peak. This indicates that during the fractionation on the DEAE-Sepharose column, either the 9/14K protein gratuitously co-eluted at the same ionic strength as SRP(S), or the protein bound to the particle, but with a substantially decreased affinity that led to its dissociation on the sucrose gradient. The small amount of 54K protein that dissociated from SRP(S) seems not to result from the nuclease treatment, as a similar fraction of the 54K protein sediments at this position when intact labelled SRP is fractionated on a sucrose gradient. The 19K protein shows the same relative intensity compared with the higher-*M_r* polypeptides as in the starting material (data not shown, but a similarly labelled SRP can be seen in Fig. 1 of ref. 13), indicating that the 19K protein remains bound stoichiometrically. This is seen much more clearly in **f**, where the peak fraction from the sucrose gradient shown in **d** is visualized by SDS-PAGE followed by staining with Coomassie blue. No protein was detected in the 1,000 mM salt eluate from the DEAE-Sepharose column, which is the step in which the *Alu*-like RNA elutes (**b**, **c**). It is possible, however, that the RNA-protein complex is unstable at the high salt concentrations used in the elution procedure. Indeed, when the peak from the 800 mM salt step was subjected to sucrose gradient analysis (data not shown), the 9/14K protein co-sedimented with the fragments of *Alu*-like RNA in that fraction. The RNA from SRP(S) was isolated by phenol/chloroform extraction and ethanol precipitation. Its 5' end was labelled with [γ -³²P]ATP using T₄ polynucleotide kinase and its length determined on polyacrylamide sequencing gels (data not shown). The major species were 149 and 150 nucleotides long, with minor species of 147 and 148 nucleotides. Using an oligonucleotide complementary to a sequence within the S segment, the precise 5' end of the fragment was determined by primer extension (data not shown). It occurs predominantly at A 102, and to a minor degree at G 101 and T 103. These analyses define both boundaries of the fragment and are indicated schematically in Fig. 4; they coincide almost precisely with those of the S segment defined by boundaries of homology (see below). The primer extension experiment also confirmed that the SRP(S) preparation was not contaminated with undigested RNA. Approximate lengths (in nucleotides) of the RNA fragments are indicated beside **a** and **b**. Positions of the six SRP polypeptides are indicated by lines in **c**, **e** and **f**, and their approximate *M_s* ($\times 10^{-3}$) are given in **c**.

translocation competence of the nascent chain. Indeed, by measuring the appearance of preprolactin in a synchronized system (Fig. 3b), we have shown that SRP(S), again in marked contrast to intact SRP, causes no detectable delay in the synthesis of the pre-secretory protein when compared with a translation reaction in which no SRP was added.

In all these respects, the subparticle is phenotypically similar in activity to an SRP composed of intact 7SL RNA and lacking the 9/14K protein; SRP(–9/14) was created by reconstituting SRP from purified components and omitting the 9/14K protein, as discussed in ref. 13. Because the purified SRP(S) also no longer contained the 9/14K protein, we expected its loss of elongation-arresting activity. We were surprised, however, to find that a full half of 7SL RNA also seems dispensable.

It seemed possible that SRP(S) retained some, albeit decreased, affinity for the 9/14K protein (see Fig. 1), and that by supplementing the subparticle with an excess of this protein, we might be able to restore the elongation-arresting activity. However, in contrast to the case of SRP(–9/14), addition of even a 10-fold molar excess of the 9/14K protein to SRP(S) in

no way restored the elongation-arrest activity, as evidenced from the rate and degree of preprolactin synthesis (Fig. 3b) or from the time window within which membranes could still be added for translocation to occur (Fig. 3a).

We conclude that SRP(S) contains all the essential components involved in signal recognition and the promotion of secretory protein translocation, and that the 9/14K protein and the *Alu*-like RNA exist in a separate and separable structural domain which, although not required for SRP-stimulated protein translocation, confers elongation-arresting activity on the particle (Fig. 4). We have also shown that regions required for the tight binding of the 9/14K protein must be contained with the *Alu*-like segments of 7SL RNA (see Fig. 1 legend). Because portions of the *Alu*-like segments were degraded by the micrococcal nuclease, we were unable to test whether this domain has any activity by itself, or whether it could restore elongation activity when added 'in trans' to SRP(S).

7SL RNA has been highly conserved through evolution⁵ across the entire length of the molecule, and *Drosophila* and human 7SL RNA can be folded into phylogenetically conserved

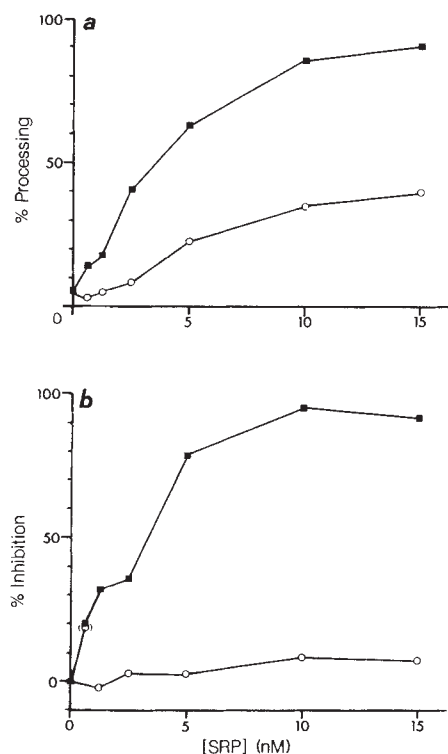


Fig. 2 Activity assays for SRP and SRP(S). *a*, Per cent processing as a function of SRP (■) or SRP(S) (○) concentration; *b*, elongation arrest as a function of SRP (■) or SRP(S) (○) concentration. SRP or SRP(S) was titrated into wheat-germ *in vitro* translations¹⁸ that were programmed with a mixture of pituitary RNA, coding primarily for the secretory protein preprolactin, and reticulocyte RNA, coding primarily for the cytoplasmic protein globin. The set of titrations shown in *a* also contained SRP-depleted microsomal membranes¹⁹ at 2 equiv. per 25- μ l translation, a concentration that had previously been shown to saturate the translocation activity (measured by per cent processing from preprolactin to prolactin) of a given concentration of SRP¹³. Translation reactions were incubated at 26 °C for 1 h and stopped by chilling on ice followed by trichloroacetic acid (TCA) precipitation. Samples were processed and per cent processing and inhibition quantitated as described elsewhere¹³.

secondary structures (see Fig. 4). This implies that 'Alu' sequences, although not present as a repetitive genomic sequence family in *Drosophila*, are part of *Drosophila* 7SL RNA. The dispensability of the 'Alu domain' *in vitro* therefore raises the question of its role *in vivo*. We have shown here that once SRP has been assembled, this portion of the particle can be removed and the subparticle remains functional. It is therefore possible that the Alu-like sequences are required for folding or stabilizing the RNA in the secondary structure necessary for proper assembly. This is plausible because in the proposed secondary structure of 7SL RNA (see Fig. 4), the Alu-like RNA portion contributes substantially to the folding energy of the molecule by forming an extended stem structure. Because Alu-like elements contain RNA polymerase III promoter function¹⁴, another possibility is that these sequences are required for the transcription of 7SL RNA. Alternatively, these sequences may be conserved because it is advantageous for the particle *in vivo* to retain elongation-arresting activity, either because arrest serves a fidelity function, preventing synthesis of preproteins in the cytoplasmic compartment, or because it can be exploited as a regulatory step, providing the cell with a means of selectively controlling secretory or membrane protein synthesis at the level of elongation. Thus far, however, elongation arrest has been documented only in the heterologous wheat-germ *in vitro* translation system. Attempts to demonstrate elongation arrest in heterologous systems com-

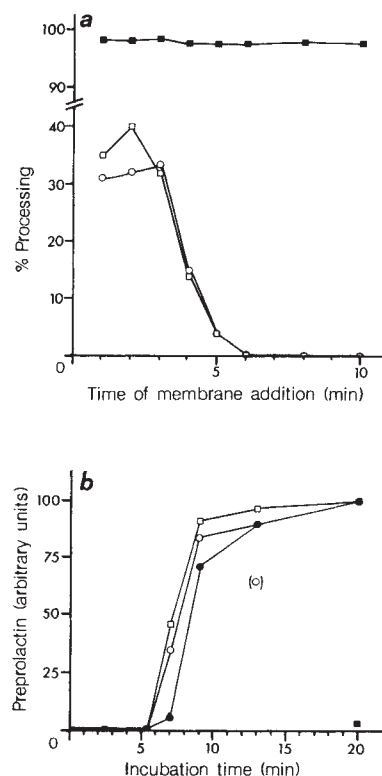


Fig. 3 Synchronized translation assays. SRP (■), SRP(S) (○), or SRP(S) supplemented with 9/14K protein (□) (see below) were added to a wheat-germ *in vitro* translation reaction to a final concentration of 14 nM. One tube contained no SRP (●). The translations were prewarmed to 26 °C for 2 min and then a mixture of pituitary and reticulocyte RNA was added to initiate protein synthesis. After 45 s, all further initiation was blocked by the addition of 7-methylguanosine 5'-monophosphate to a final concentration of 4 mM. At various times thereafter, 5- μ l aliquots were added either to 5 μ l of 20% TCA, to measure preprolactin synthesis, or to 0.5 μ l of SRP-depleted microsomal membranes, to measure translocation as a function of the time of membrane addition. In the case of the membrane addition experiments, synthesis was allowed to continue for 20 min. Samples were processed as for Fig. 2 except that gels were subjected to fluorography before autoradiography. *a* Shows per cent processing as a function of the time of membrane addition. If SRP was present throughout the experiment, membranes could be added at any time point tested and productive translocation would still occur. However, if SRP(S) was present, processing was strictly dependent on the time at which membranes were added; translocation (measured by % processing) was reduced 50% at ~3.9 min, and was completely abolished when membranes were added after 5 min of elongation. This corresponds to a stage of elongation in which approximately half of the preprolactin molecule has been polymerized (see *b*) and agrees well with the size of the arrested fragment synthesized in the presence of SRP¹². The 9/14K protein was preincubated with SRP(S) under ionic conditions used to reconstitute SRP from RNA and proteins¹³ in order to facilitate any functional rebinding of the 9/14K protein to SRP(S). Note that the addition of a 10-fold molar excess of the 9/14K protein has no effect on the time dependence of the process. *b* Shows that neither SRP(S) nor SRP(S) supplemented with 9/14K protein measurably decreased the rate of preprolactin synthesis.

prised of mammalian components have failed¹⁵. However, no mammalian system studied thus far has responded to the addition of exogenous SRP either by elongation arrest or by increased levels of translocation; elongation arrest in these systems may become apparent once they are dependent on exogenously added SRP. Ultimately, however, due to the *in vitro* nature of these approaches, the relevance of elongation arrest still remains to be demonstrated *in vivo*.

Alu-like sequences are present in other cytoplasmic RNAs^{1,3}

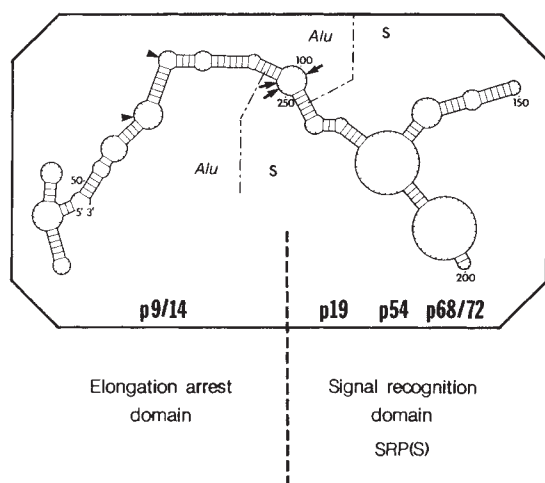


Fig. 4 SRP is composed of two separable domains. A possible secondary structure for human 7SL RNA is shown (similar secondary structures have been proposed by Gundelfinger *et al.*²⁰, E. Ullu (personal communication) and C. Zwieb²¹). Connecting lines indicate base pairs; G-U base pairs are included. We have sequenced canine 7SL RNA by a combination of RNA sequencing⁴ and primer extension in the presence of dideoxy nucleotides (data not shown). The sequence differs from human 7SL RNA at only four positions. The nucleotide present in the canine RNA at these positions is as follows: C 19, A 53, C 69 and C 87. None of these differences alters the predicted secondary structure. Micrococcal nuclease cleaves the particle at the arrows, removing the elongation arrest-promoting domain, which is comprised of the *Alu*-like RNA and the 9/14K protein (the thin dashed lines indicate the boundaries of homology^{5,7}). SRP(S) retains signal recognition and protein translocation-promoting functions. The major 5' and 3' ends of the S fragment are indicated by arrows. Additional cuts by micrococcal nuclease at positions determined by Gundelfinger *et al.*¹¹ yield the RNA fragments shown in Fig. 1a and are indicated by arrowheads.

(for example the 3'-untranslated region of the low-density lipoprotein (LDL) receptor mRNA¹⁶). It is possible that these sequences form similar structural domains that could bind the 9/14K or related proteins. Indeed, two of the *Alu* sequences in the LDL receptor mRNA can be folded into secondary structures closely resembling the *Alu* 'end' of 7SL RNA (P.W., unpublished). If these domains are functional, then, in analogy to their role in SRP, *Alu* transcripts that are transported to the cytoplasmic compartment may have some role in translational control.

We thank P. Hoben for helpful discussions. This work was supported by NIH grant GM-32384. P.W. is a recipient of support from the Chicago Community Trust/Searle Scholars Program.

Received 16 September; accepted 19 December 1985.

- Schmid, C. W. & Jelinek, W. R. *Science* **216**, 1065-1070 (1982).
- Sharp, P. A. *Nature* **301**, 471-472 (1983).
- Calabretta, B., Robberson, D. L., Maizel, A. L. & Saunders, G. F. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6003-6007 (1981).
- Walter, P. & Blobel, G. *Nature* **299**, 691-698 (1982).
- Ullu, E. & Tschudi, C. *Nature* **312**, 171-172 (1984).
- Walter, P., Gilmore, R. & Blobel, G. *Cell* **38**, 5-8 (1984).
- Ullu, E., Murphy, S. & Melli, M. *Cell* **29**, 195-202 (1982).
- Andrews, D. W., Walter, P. & Ottensmeyer, F. P. *Proc. natn. Acad. Sci. U.S.A.* **82**, 785-789 (1985).
- Walter, P. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7112-7116 (1980).
- Walter, P. & Blobel, G. *Cell* **34**, 525-533 (1983).
- Gundelfinger, E. D., Krause, E., Melli, M. & Dobberstein, B. *Nucleic Acids Res.* **11**, 7363-7373 (1983).
- Walter, P. & Blobel, G. *J. Cell Biol.* **91**, 557-561 (1981).
- Siegel, V. & Walter, P. *J. Cell Biol.* **100**, 1913-1921 (1985).
- Sakonju, S., Bogenhagen, D. F. & Brown, D. D. *Cell* **19**, 13-25 (1980).
- Meyer, D. I. *EMBO J.* **4**, 2031-2033 (1985).
- Yamamoto, T. *et al. Cell* **39**, 27-38 (1984).
- Walter, P. & Blobel, G. *Meth. Enzym.* **96**, 682-691 (1983).
- Erickson, A. H. & Blobel, G. *Meth. Enzym.* **96**, 38-50 (1983).
- Walter, P. & Blobel, G. *Meth. Enzym.* **96**, 84-93 (1983).
- Gundelfinger, E. D., Carlo, M. D., Zopf, D. & Melli, M. *EMBO J.* **3**, 2325-2332 (1984).
- Zwieb, C. *Nucleic Acids Res.* **13**, 6105-6124 (1985).

Localization of gene for human p53 tumour antigen to band 17p13

M. Isobe*, B. S. Emanuel†, D. Givol‡, M. Oren‡ & C. M. Croce*

* Wistar Institute, 3601 Spruce Street and † Departments of Pediatrics and Human Genetics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA

‡ Department of Chemical Immunology, Rehovot 76100, Israel

Recently the gene for the cellular tumour antigen p53, a phosphoprotein found in increased concentration in a variety of human cells¹, has been mapped to region 17q22 by *in situ* hybridization techniques and has been shown to translocate to the chromosome carrying the translocation [t(15; 17)] associated with acute promyelocytic leukaemia (APL)². Based on this finding it has been postulated that this gene has a role in the pathogenesis of APL². Here we present evidence that the gene for p53 is not located on the long arm of chromosome 17, but maps to band 17p13. We therefore suggest that this gene is not directly involved in the chromosome translocation observed in APL.

It has been suggested that p53 is a protein related to the cell cycle and that it may be involved in the control of cell proliferation³⁻⁵. By introducing the p53 gene into Abelson murine leukaemia virus-transformed cells that lack p53, it was shown that overproduction of p53 rendered cells malignant⁶. The expression of a completely transformed phenotype requires the concomitant expression of *abl* and p53 (ref. 6). Moreover, primary embryonic fibroblasts can be transformed by co-transfection of the cells with *ras* and p53 genes⁷⁻⁹, suggesting that p53 belongs to the nuclear oncogene family.

It has been demonstrated that the mouse genome contains two p53-specific genes¹⁰. By using somatic cell hybrids, the functional p53 gene was mapped to mouse chromosome 11, while the pseudogene was shown to reside on chromosome 14 (refs 11, 12). Thus, it would be of interest to determine the chromosomal location of the gene in the human genome as specific translocations are often associated with altered expression or modification of normal proto-oncogenes. Moreover, putative oncogenes have been identified by virtue of their presence at the junction between chromosomes involved in specific translocations in leukaemia and lymphoma^{13,14}.

To determine whether the p53 gene has a role in APL, we examined somatic cell hybrids containing different regions of chromosome 17 for the presence of the human p53 gene. We found that hybrid GM54VA × BALB/cIIa Cl31, which contains only human chromosome 17 and no other human chromosome¹⁵, contains the p53 gene as well as other markers of chromosome 17. This result is consistent with previous reports which assigned the p53 gene to chromosome 17 (ref. 16) (Table 1, Fig. 1). We

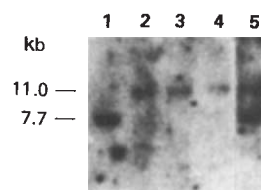


Fig. 1 Southern blotting analysis of hybrid cell DNA after digestion with *Bam*HI and hybridization with a human p53 complementary DNA probe. Lane 1, human DNA; lane 2, mouse DNA; lane 3, hybrid N9, which contains only the long arm of chromosome 17 (ref. 15); lane 4, hybrid 275S, which contains only the region 17q22-qter translocated to chromosome 21 (ref. 15); lane 5, GM54VA × BALB/cIIa Cl31 DNA (hybrid with only human chromosome 17). The probe was a full-length p53 cDNA (pHP53) in pBR322 (ref. 17). kb, kilobases.

Fig. 2 Localization of the gene for p53 in the human genome by chromosomal *in situ* hybridization. *a*, Autoradiograph of a representative G-banded metaphase spread hybridized with the human p53 probe. The arrow indicates a grain on the short arm of chromosome 17. *b*, Diagram showing the grain distribution in 80 metaphase spreads. The abscissa represents the chromosomes in their relative size proportion. The distribution of 179 grains on 80 spreads was scored; 45 were found overlying 17p11.2-17pter. The p53 cDNA clone, ph53, is a full-length cDNA of 2.5 kb inserted in pBR322 at the *Bam*HI site¹⁷.

Methods. Metaphase spreads were prepared from normal human male lymphocyte cultures which had been stimulated with phytohaemagglutinin for 72 h *in vitro*. p53 was labelled by nick-translation with ³H-dCTP (62 Ci mmol⁻¹; 1 Ci = 37 GBq; NEN), ³H-dGTP (39.9 Ci mmol⁻¹), ³H-dTTP (100.1 Ci mmol⁻¹) and ³H-dATP (51.9 Ci mmol⁻¹). The techniques used for *in situ* hybridization were essentially as described by Harper and Saunders¹⁸. Chromosome preparations were treated with 100 µg ml⁻¹ RNase A (Sigma) in 2×SSC for 20 min at 37 °C, then immersed in 70% formamide in 2×SSC, pH 7.0 at 70 °C for 2 min to denature the chromosomal DNA. The chromosome preparations were then hybridized with ³H-labelled p53 probe (specific activity 1.1 × 10⁷ c.p.m. µg⁻¹) at a concentration of 70 ng ml⁻¹ in 50% formamide, 2×SSCP (150 mM NaCl, 15 mM sodium citrate, 20 mM NaH₂PO₄, 10% dextran sulphate; Pharmacia), pH 7.0 for 20 h at 37 °C. A 1,000-fold excess of sonicated salmon sperm DNA was included as carrier. Slides were rinsed thoroughly in 50% formamide, 2×SSC at 39 °C, rinsed thoroughly in 2×SSC at 39 °C and then dehydrated in ethanol. Hybridized slides were coated with Kodak NTB2 emulsion (diluted 1:1 with water), and exposed at 4 °C for 10–16 days. Autoradiographs were developed with Kodak Dektol at 15 °C. The chromosomes were then G-banded essentially as described by Cannizzaro and Emanuel¹⁹, with a mixture of 6 parts borate buffer (50 mM Na₂SO₄/2.5 mM Na₂B₄O₇, pH 9.2) to 1 part Wright's/Giemsa stain solution (2.5 g of Wright's stain per l/1.4 g of Giemsa stain per l in methanol).

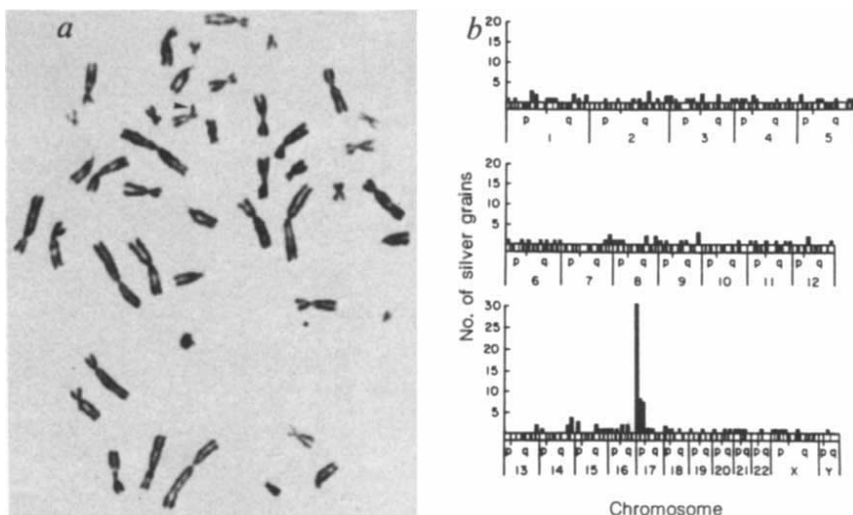


Table 1 Presence of chromosome 17 markers in mouse-human hybrids

Hybrid	Human DNA markers			Human isozymes		Chromosome 17 present in hybrid
	<i>c-erb-A</i>	NGFR	p53	<i>tk</i>	<i>gk</i>	
GM54VA × BALB/cIIa C131	+	+	+	+	+	17
N9	+	+	—	+	+	17q
275S	—	+	—	+	+	17q22-qter

We have previously mapped the *c-erb-A* oncogene to region 17q21-q22, proximal to the 17q22 breakpoint observed in APL¹⁶. We have also mapped the gene for the nerve growth factor receptor (NGRF) to band 17q22, distal to the 17q22 breakpoint observed in APL¹⁵. The genes for thymidine kinase (*tk*) and galactokinase (*gk*) are distal to the breakpoint observed in APL^{15,16}.

found, however, that hybrid N9, which contains only human chromosomes 6, 7, 21 and the long arm of chromosome 17 (17q)¹⁵, does not contain the human p53 gene (Fig. 1, lane 3; Table 1). Similarly, hybrid 275S, which carries the segment 17q22-qter translocated to chromosome 21 (refs 15, 16) and has lost all other human chromosomes, is negative for the p53 gene (Fig. 1, lane 4; Table 1). These results indicate that the gene for human p53 is not located on the long arm of chromosome 17 as reported previously² and therefore it cannot be involved in APL.

To confirm and refine the somatic cell hybrid results, *in situ* hybridization to human metaphase chromosomes was performed using the previously described human p53 plasmid¹⁷. After autoradiography, metaphase spreads were analysed for grain localization; an example of one such metaphase spread is shown in Fig. 2a. Over 25% of all grains were located on the short arm of chromosome 17, and over 97% of the 17p grains were between 17p11.2 and 17pter, with most grains at 17p13. Figure 2b shows a histogram of the silver grain distribution on all the chromosomes. The short arm of chromosome 17 represents ~1.0% of the haploid genome, and our observation that more than 25% of the human p53 probe hybridization was localized to this region is highly significant ($P < 0.001$). Thus, cytological hybridization sublocalizes the human p53 gene to the region between 17p11.2 and 17pter.

As the gene for p53 is located on the short arm and not on the long arm of chromosome 17, it seems highly unlikely that this gene is involved in the pathogenesis of APL as proposed by others².

This work was supported by NIH grants CA-39860 and GM32592.

Received 3 October 1985; accepted 21 January 1986.

- Klein, G. (ed.) *Advances in Viral Oncology* Vol. 2 (Raven, New York, 1982).
- LeBeau, M. M., Westbrook, C. A., Diaz, M. O., Rowley, J. D. & Oren, M. *Nature* **316**, 826–828 (1985).
- Milner, J. & Milner, S. *Virology* **112**, 785–788 (1981).
- Mercer, W. E., Nelson, D., DeLeo, A. B., Old, L. J. & Baserga, R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6309–6312 (1982).
- Reich, N. C. & Levine, A. J. *Nature* **308**, 199–201 (1984).
- Wolf, D., Harris, N. & Rotter, V. *Cell* **38**, 119–126 (1984).
- Eliyah, D., Raz, A., Gruss, P., Givol, D. & Oren, M. *Nature* **312**, 646–649 (1984).
- Parada, L. F., Land, H., Weinberg, R. A., Wold, D. & Rotter, V. *Nature* **312**, 649–651 (1984).
- Jenkins, J. R., Rudge, K. & Currie, G. A. *Nature* **312**, 651–654 (1984).
- Oren, M., Bienz, B., Givol, D., Rechavi, G. & Zakut, R. *EMBO J.* **2**, 1633–1639 (1983).
- Rotter, V., Wolf, D., Pravtcheva, D. & Ruddle, F. H. *Molec. cell. Biol.* **4**, 383–385 (1984).
- Czosnek, H. H. *et al. Molec. cell Biol.* **4**, 1638–1640 (1984).
- Tsujimoto, Y., Yunis, J., Onorato-Showe, L., Nowell, P. C. & Croce, C. M. *Science* **224**, 1403–1406 (1984).
- Tsujimoto, Y., Finger, L., Yunis, J., Nowell, P. C. & Croce, C. M. *Science* **226**, 1097–1099 (1984).
- Huebner, K. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Dayton, A. I. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 4495–4499 (1984).
- Zakut-Houri, R., Bienz-Tadmor, B., Givol, D. & Oren, M. *EMBO J.* **4**, 1251–1255 (1985).
- Harper, M. E. & Saunders, G. F. *Chromosoma* **83**, 431–439 (1981).
- Cannizzaro, L. A. & Emanuel, B. S. *Cytogenet. Cell Genet.* **38**, 308–309 (1984).

Mesozoic and Cenozoic evolution of the North and South China blocks

RECENT palaeomagnetic studies have focused on the deformation history of eastern Asia in relation to the north-western Pacific subduction zones and the collision of India¹⁻⁴. For this, reliable apparent polar wander (APW) paths for major cratonic blocks, such as those of north and south China (respectively NCB and SCB), India and Eurasia, are essential, and the recent work of Lin *et al.*^{5,6} is a major contribution to the reference APW paths of the NCB and SCB. These new data, believed to demonstrate a 4,000-km eastward displacement of the China blocks with respect to northern Eurasia⁵, are actually consistent with other scenarios, some of which are easier to reconcile with the local geology.

Our Fig. 1 shows the Middle Jurassic to Upper Cretaceous poles of both the SCB and NCB as given in ref. 5 (Table 1). As noted by Lin *et al.*, the Middle Jurassic poles from both blocks are statistically indistinguishable; these poles are within the common intersection of their 95% domains of confidence. The same is true respectively of the pairs of Upper Jurassic, Lower Cretaceous and Upper Cretaceous poles from the NCB and SCB (note that the Upper Cretaceous poles do not come from mainland China but from northern and southern Korea, respectively; these poles appear to require revision and will therefore not be used here; see ref. 7). It appears, therefore, that the sites sampled on the NCB and SCB belonged to a single stable unit from the Middle Jurassic onwards. Moreover, all the poles fall close to each other and no significant trend emerges from Fig. 1. We do not think that the loops and details of the APW paths for the Mesozoic shown in Fig. 2 of the study by Lin *et al.*⁵ can be resolved with available data. We can compute a mean Middle Jurassic to Lower Cretaceous pole for the NCB (74° N, 214° E; $A_{95} = 6.5^\circ$, where A_{95} is the 95% error in Fisher's⁸ analyses) and SCB (73° N, 212° E; $A_{95} = 7.5^\circ$). These are virtually identical and can in turn be combined to provide a Middle Jurassic to Lower Cretaceous China pole at 74° N, 214° E ($A_{95} = 4^\circ$; pole C in Fig. 2). Thus, China (as represented by sites in Shandong and Zhejiang provinces only) stood at constant latitudes, essentially similar to present latitudes, from the Middle Jurassic to the Lower Cretaceous, and probably up to the present (see also ref. 7).

These are not the latitudes that would have been expected had China already occupied its present position with respect to northern Eurasia⁵⁻⁷. Irving's⁹ APW path

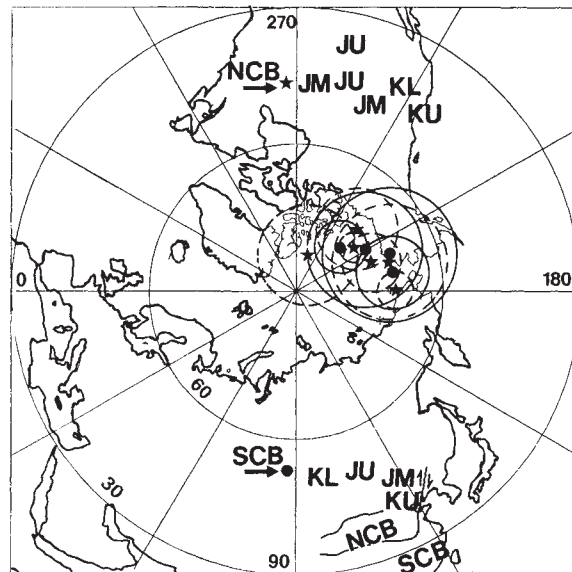
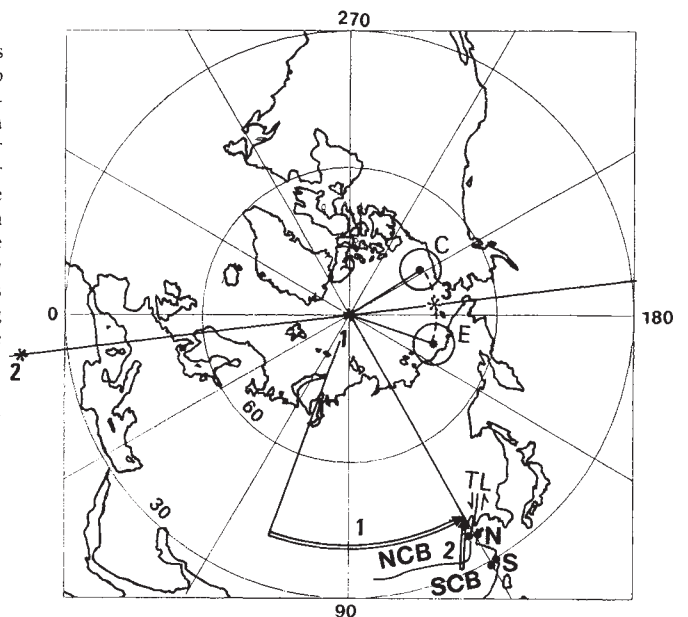


Fig. 1 Middle Jurassic to Upper Cretaceous poles for the North China (NCB) and South China (SCB) blocks (from ref. 5). Poles from the SCB are shown as dots and those from the NCB as stars, surrounded by their A_{95} confidence circles. The ages of each pole are given in sequence on the sides of the figure to avoid confusion. KU, Upper Cretaceous; KL, Lower Cretaceous; JU, Upper Jurassic; JM, Middle Jurassic.

Fig. 2 Mean poles (Middle Jurassic to Lower Cretaceous, 170–100 Myr), C for China (this paper) and E for northern Eurasia (after ref. 9). Also shown are the sampling sites N in the NCB and S in the SCB⁶, the boundary following the Qinling mountains and the Tancheng-Lujian (TL) fault system, and three possible Euler poles that will bring poles E and C in coincidence: (1) is from Lin *et al.*⁵, (2) is the pole proposed in this paper and (3) is another, geologically unreasonable, pole, all of them lying on the 7° E–187° E meridian. Corresponding motions (1) and (2) for a reference point on the NCB are shown as double arrows.



for northern Eurasia predicts a fairly stable pole at 73° N, 160° E ($A_{95} = 4^\circ$; pole E in Fig. 2) for the period of interest here (Middle Jurassic–Lower Cretaceous; that is 170–100 Myr). This is significantly different from the China pole (Fig. 2) and would seem to imply post-Lower Cretaceous displacement between the two land masses. Since the two poles lie at essentially the same latitude, they can be reconciled by a $\sim 50^\circ$ rotation around the present geographic North Pole (Fig. 2, rotation pole 1). This is in essence the interpretation of Lin *et al.*⁵, which implies $\sim 4,000$ km of eastward motion of the SCB with respect to stable northern Eurasia

(Fig. 2, double arrow labelled 1). Lin *et al.* propose that at least 2,500 km of that displacement occurred after the collision of India along the Central Asian fold belt, which marks the boundary of the NCB and northern Eurasia. This very large Cenozoic motion seems to contradict available geological evidence. In a detailed study of active faulting and Cenozoic tectonics in China, Tien Shan, Mongolia and Baikal regions, Tapponnier and Molnar¹⁰ estimated left-lateral offsets of no more than 100 km on no more than three major faults (including the Gobi-Altai and Bolnai-Khangai faults). Thus, the total left-lateral (eastwards)

Cenozoic motion of China with respect to stable Eurasia is unlikely to exceed 500 km.

Actually, the rotation around the North Pole proposed by Lin *et al.* is only one of an infinity of rotations that can bring the China and northern Eurasia Middle Jurassic/Lower Cretaceous poles in coincidence: indeed, all Euler poles falling along the great circle that bisects the two mean poles are acceptable (Fig. 2). Another solution, implying the smallest possible rotation ($13 \pm 5^\circ$), has a pole lying at 16° N, 7° E (rotation pole 2 in Fig. 2). For the Chinese sites located near 40° N, 120° E, this pole implies 'only' 1,300 km of displacement with respect to northern Eurasia, with an azimuth of $\sim 30^\circ$ E (Fig. 2, double arrow labelled 2). Of course, any pole falling between the Lin *et al.*⁵ Euler pole 1 and the pole that leads to the minimum displacement (pole 2) is acceptable from a palaeomagnetic (if not geological) point of view.

If one accepts a Euler pole close to pole 2 as a reasonable solution, where can the 1,300-km displacement occur? Lee *et al.*⁷ point out that all available palaeomagnetic sites from the China blocks come from close to or east of the Tancheng-Lujiang (Tanlu) fault zone, which is known to have undergone major sinistral slip during the Mesozoic¹¹. All the Upper Jurassic to Lower Cretaceous sites given by Lin⁶ fall within two very small (100–150 km wide) areas, respectively near Hangzhou in Zhejiang province and Jinan in Shandong province (points S and N in Fig. 2). Thus, it is not obvious that available Jurassic and Cretaceous Chinese data pertain to the bulk of the NCB, west of the Tanlu fault; this part of the block (Ordos and Tarim, that is, the largest part of it) is still devoid of palaeomagnetic sites and thus is a primary goal for future sampling.

In a more detailed discussion of Cretaceous data from China and Korea, Lee *et al.*⁷ point out that the remaining latitudinal discrepancy between China and Eurasia could be accounted for in at least three ways: (1) The post-Lower Cretaceous motion could have occurred between the SCB and NCB in the Qinling mountains and along the Tanlu fault. However, motions documented by geology^{11,12} and a palaeomagnetism^{5,6,13} would seem to have taken place earlier. (2) The motion could have occurred between the NCB and Siberia in the Central Asian fold belt. But we have already pointed out that such major belated motion was most unlikely¹⁰. (3) The palaeomagnetic data for Siberia (Cretaceous Eurasia) are not reliable and should be revised⁷. It is actually our feeling that extensive work on the Cretaceous APW path of Eurasia, and particularly Siberia, is urgently needed.

This is IGP contribution NS 891. We thank J. Lin and M. Fuller for comments

and for making a copy of ref. 6 available to us.

V. COURTILLOT

J. BESSE

*Institut de Physique du Globe,
Laboratoire de Géomagnétisme et
Paléomagnétisme (UA CNRS 729),
Science Physique de la Terre,
Université Paris VII,
75232 Paris, Cedex 05, France*

1. Achache, J., Courtillot, V. & Besse, J. *Earth planet. Sci. Lett.* **63**, 123–136 (1983).
2. Klootwijk, C. T. & Radhakrishnamurty, C. in *Paleoreconstructions of the Continents* (eds McElhinny, M. W. & Valencio, P. A.) 93–105 (American Geophysical Union, Washington, DC, 1981).
3. Achache, J., Courtillot, V. & Zhou, Y. *J. geophys. Res.* **89**, 10311–10339 (1984).
4. Besse, J., Courtillot, V., Pozzi, J. P., Westphal, M. & Zhou, Y. *Nature* **311**, 621–626 (1984).
5. Lin, J., Fuller, M. & Zhang, W. *Nature* **313**, 444–449 (1985).
6. Lin, J. thesis, Univ. California (1984).
7. Lee, G., Besse, J. & Courtillot, V. *J. geophys. Res.* (submitted).
8. Fisher, R. A. *Proc. R. Soc. A* **217**, 295–305 (1953).
9. Irving, E. *Nature* **270**, 304–309 (1977).
10. Tapponnier, P. & Molnar, P. *J. geophys. Res.* **82**, 2905–2930 (1977); **84**, 3425–3459 (1979).
11. Xu, J. *Int. geol. Congr.* **26**, (abstr.) 1, 410 (1980).
12. Mattauer, M. *et al. Nature* **317**, 496–500 (1985).
13. Cheng, G. & Fang, Z. *Seismology and Geology* Vol. 2, 4 (Peking, 1980).

LIN AND FULLER REPLY—Courtillot and Besse raise two major questions about our paper¹: the uniqueness of our Euler pole and the offset along the Tanlu fault in eastern China.

In answer to their first question, we emphasize that our Euler pole is unique. Although a unique Euler pole cannot be obtained if it is based on only one pair of palaeopoles, it certainly can be achieved if one has several pairs of palaeopoles, or better still, two pieces of curves. Our Euler pole (86.5° N, 31.5° E) was obtained by matching two pieces of curves, the late Palaeozoic–Mesozoic apparent polar wander (APW) paths for the South China block (SCB) and for northern Eurasia, each of which consists of at least nine palaeopoles. It is on the basis of this unique Euler pole that we suggest the SCB moved $4,300 \pm 1,200$ km eastwards with respect to northern Eurasia.

As to the Tanlu fault in eastern China, most Chinese geologists agree that it underwent an early Cretaceous sinistral offset, but no consensus has been reached over the amount of the offset; most people suggest that it was about 200–300 km. The largest offset ever suggested², based on an analysis of the regional geology, is about 600 km. The 1,300-km displacement suggested by Courtillot and Besse is more than double this. A palinspastic reconstruction would place the Shandong and Liaodong peninsulas of North China in the present location of Fujian and Guangdong provinces of South China, or even in the South China Sea if we accept Courtillot and Besse's suggestion. Most Chinese geologists would find it difficult to accept such a palaeogeographic configuration.

Recent new Triassic and Jurassic palaeomagnetic results, despite the fact that they come from Guizhou and Sichuan provinces of southwestern China, several thousand kilometres away from the Tanlu fault^{3,4}, are consistent with our SCB APW path. This indicates that our sampling area in Zhejiang has been a coherent part of the SCB since the early Mesozoic, maybe even since the early Palaeozoic, as we argued earlier by comparing the Cambrian palaeomagnetic results from Zhejiang, Hubei and Yunnan provinces^{1,5}.

JIN-LU LIN

*Institute of Geology,
Academia Sinica,
Beijing, China*

M. FULLER

*Department of Geological Sciences,
University of California,
Santa Barbara,
California 93106, USA*

1. Lin, Jin-lu, Fuller, M. & Zhang, Wen-you *Nature* **313**, 444–449 (1985).
2. Xu, Jia-wei *Int. geol. Congr.* **26**, (abstr.) 1, 410 (1980).
3. Opdyke, N. D., Huang, K., Kent, D. V., Xu, G. & Zhang, W. *Trans. Am. geophys. Un.* **65**(45), 862 (1984).
4. Zhu, Zhi-wen, Hao, Tien-yao & Zhao, Hui-sheng in *Int. Symp. Deep Internal Processes and Continental Rifting (DIPCR)*, China, (abstr.) 248 (Chengdu, 1985).
5. Lin, Jin-lu, Fuller, M. & Zhang, Wen-you *J. Geodyn.* **2**, 91–114 (1985).

Liposomal subunit vaccine against Epstein-Barr virus-induced malignant lymphoma

THE recent success of Epstein *et al.*¹ in protecting cottontop tamarins against Epstein-Barr virus-induced malignant lymphoma by using the purified virus-determined membrane antigen gp340 incorporated in liposomes, has once more highlighted the need for an effective immunological adjuvant. The antigen gp340 is poorly immunogenic but in its liposomal form elicited specific high titres of virus-neutralizing antibodies in mice^{2,3}. However, in tamarins the production of titres sufficiently high to confer protection¹ was slow and required multiple (intraperitoneal) injections. Epstein *et al.*¹ refer to more recent work⁴ with tamarins in which larger doses of liposomal gp340 produced rapid induction of high-titre antibodies. As discussed previously^{1,5}, the vaccine could be improved upon by presenting it in such a manner that a protective immune response is achieved with lesser amounts and fewer doses of gp340 and by altering the schedule of immunization. In this respect, a role for liposomes in immunopotentiality appears in the paper¹ to be somewhat less promising than it actually is.

Liposomes⁶ are uniquely versatile in size, surface characteristics, lipid composition and in the ways in which they can accommodate antigens, preparations now

being available in a freeze-dried form⁷. Since 1974 a large number of antigens relevant to human and veterinary immunization have been incorporated in liposomes⁸: the list⁶ includes diphtheria toxoid, *Plasmodium falciparum* antigens, synthetic peptides for hepatitis B and foot-and-mouth disease vaccines, cholera toxin and *Streptococcus mutans* cell-wall antigens. Studies⁶ in numerous laboratories have shown that optimal immunization with a given antigen can be achieved if the structural characteristics of liposomes, the way in which the antigen is accommodated (if needed together with other immunopotentiating agents) and the route and schedule of administration are judiciously chosen. Also, targeting⁶ of antigens to immunocompetent cells by means of liposomes coated with cell-specific ligands is likely to improve immunogenicity considerably.

It follows that the type of liposomes used by Epstein *et al.*¹ may have contributed to the poor immune response to gp340 in the animal species used. As the liposomes were made^{2,3} of L- α -phosphatidylcholine only, they would be expected, on reaching⁶ the circulating blood from the peritoneal cavity, to rapidly disintegrate through the action of high-density lipoproteins which remove phospholipid molecules from vesicles devoid of cholesterol⁶. Thus, only a fraction of the original dose of liposomes still containing the antigen would have reacted with immunocompetent cells. Liposomal size also controls vesicle stability as well as uptake by tissues, including lymph nodes⁶. There are other variables⁶ which affect liposomal adjuvanticity (for example, membrane fluidity and surface charge), but of particular interest is the finding⁹ that covalent attachment of antigens on the liposomal surface further augments immunogenicity and this may be applicable to gp340. It is unfortunate that the limited number of tamarins available has prevented Epstein *et al.*¹ from evaluating the many possible versions of their liposomal vaccine.

GREGORY GREGORIADIS

Medical Research Council Group,
Academic Department of Medicine,
Royal Free Hospital School
of Medicine, Pond Street,
London NW3 2QG, UK

Life history correlations and demography

IN a recent paper, Harvey and Zammuto¹ pointed out that variation in age at first reproduction among female mammals is strongly correlated with expectation of life at birth, even after the effects of body size have been removed. We wish to make three points about the interpretation of that correlation. The second and third are of wider interest as they are relevant to a number of other studies that present relationships between life history variables.

First, expectation of life at birth is the sum of two components, pre-reproductive lifespan and reproductive lifespan. The correlation is therefore between pre-reproductive lifespan (age at first reproduction) and the sum of the two parts (expectation of life at birth). The same is true of a similar plot for iteroparous plant species given by Harper (ref. 2, page 688) in which age at first reproduction is plotted against lifespan. It is not surprising that the sum is correlated with one of its parts, irrespective of effects of body size. This complication can be removed by correlating age at first reproduction with reproductive lifespan rather than total lifespan. For Harvey and Zammuto's case¹, the correlation is highly significant even with the effects of body size removed by partial correlation ($r = 0.76$, $n = 25$, $P < 0.001$).

Our second point follows because in a constant-sized population, natality must be balanced by mortality. This makes correlations between at least some components of life history inevitable, and a zero correlation between any two components *a priori* unlikely. In the above case, change in life expectation could be balanced by variation in litter size, inter-birth interval or age at first reproduction. In fact, as Harvey and Zammuto pointed out, age at first breeding seemed to be the most relevant variable.

To make this point formally, consider a sexual population of constant size and age structure, and sexual symmetry. Let π be the probability of surviving to reproduce, μ the mean adult mortality per year, L the expected reproductive lifespan and M the average number of offspring produced by a female chosen at random from those reproducing in a given year. Then

$$\pi LM = \pi M / \mu = 2 \quad (1)$$

This relationship could explain the results of a number of studies. Stearns³ points out a relationship between the ratio of juvenile to adult mortality ($= (1 - \pi)L$) and number of breeding seasons (L), while Tinkle⁴ and O'Connor⁵ report one between fecundity (M) and survivorship ($1 - \mu$). The correlation reported above might also be explained in the same way because it is between reproductive lifespan (L) and age at first reproduction, which may be related inversely to the probability

of survival to first breeding (π). This case is more tenuous as it involves another relationship in addition to equation (1). Harper's² correlation for iteroparous plant species (see above) may also be of this kind.

Unless other variables adjust to compensate, the relationships reported are in each case an inevitable demographic consequence of a net growth rate of zero. When variables are related by an exact formula such as equation (1), it is barely possible for each pair of variables to be uncorrelated. Knowing that species differ in longevity, for example, therefore makes it inevitable that some pairs of variables will be correlated.

Our final point is that it is useful to distinguish between two kinds of explanation for any pattern that holds between life history variables: adaptive and non-adaptive. The distinction was implied but not made explicit in the final paragraph of Harvey and Zammuto's paper¹. They pointed out that, among Columbian ground squirrels, increase in food supplies results in earlier reproduction and higher mortality. Harvey and Zammuto considered these populations to differ in age at first breeding and reproductive lifespan for non-adaptive reasons, while interspecies differences were attributed to adaptive reasons. After showing that, irrespective of size, genera with later ages at first reproduction had lower mortality rates, they stated that "this calls for an explanation in terms of the costs and benefits of reproduction at different ages". Their reason for making this claim was that the correlation seemed to imply that some genera delay reproduction beyond the time when they become physically capable of bearing viable young. Only further work can determine whether this is so.

Life histories are often seen as strategies of animals. They may be more profitably viewed as consequences of animals' actions (which may be evolved strategies), environmental effects and demographic constraints.

We thank Andrew Watkinson and Mark Ridley for useful comments.

WILLIAM J. SUTHERLAND

School of Biological Sciences
University of East Anglia,
Norwich NR4 7TJ, UK

ALAN GRAFEN

PAUL H. HARVEY

Department of Zoology,
University of Oxford,
South Parks Road,
Oxford OX1 3PS, UK

1. Epstein, M. A., Morgan, A. J., Finerty, S., Randle, B. J. & Kirkwood, J. K. *Nature* **318**, 287-289 (1985).
2. Morgan, A. J., Epstein, M. A. & North, J. R. *J. med. Virol.* **13**, 281-292 (1984).
3. North, J. R., Morgan, A. J., Thompson, J. L. & Epstein, M. A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7504-7508 (1982).
4. Epstein, M. A., Randle, B. J., Finerty, S. & Kirkwood, J. K. *Clin. exp. Immun.* (in the press).
5. Beale, A. J. *Nature* **318**, 230 (1985).
6. Gregoriadis, G. *Trends Biotechnol.* **3**, 235-241 (1985).
7. Kirby, C. & Gregoriadis, G. *Biotechnology* **2**(11), 979-984 (1984).
8. Allison, A. C. & Gregoriadis, G. *Nature* **252**, 252 (1974).
9. Snyder, S. L. & Vannier, W. E. *Biochim. biophys. Acta* **772**, 288-294 (1984).

1. Harvey, P. H. & Zammuto, R. M. *Nature* **315**, 319-320 (1985).
2. Harper, J. L. *Population Biology of Plants* (Academic, London, 1977).
3. Stearns, S. C. *Q. Rev. Biol.* **51**, 3-47 (1976).
4. Tinkle, D. W. *Am. Nat.* **103**, 501-516 (1979).
5. O'Connor, R. J. in *Behavioural Ecology, Ecological Consequences of Adaptive Behaviour* (eds Sibley, R. M. & Smith, R. H.) 105-142 (Blackwell, Oxford, 1985).

Laboratory automation with robotics

from James N. Little

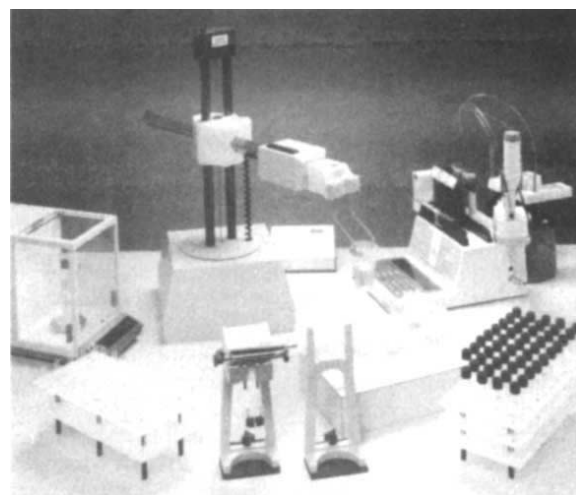
Robotics technology is destined to have a major impact on operations and procedures in tomorrow's laboratories. It should put analysis in the fast lane.

New technology using laboratory computers for analytical measurement and for data reduction has developed at a rapid rate, but sample preparation technology has not kept pace. The great emphasis on instrumental measurement techniques such as chromatography and spectroscopy often obscures the importance of sample handling and preparation in achieving quality analytical results. Today, sample preparation is the weak element in analytical methods. It is the area that needs improvement if total laboratory automation is to be achieved.

Most chemical samples require several preparation steps prior to instrumental analysis and, with the exception of automatic sample injection, most procedures continue to be performed much as they have in the past. Subtle differences between analytes, interferences and matrices give rise to a great variety of sample preparation methods, so automation instruments have to be flexible if they are to be generally useful. Laboratory robotics deals with the variety of preparation requirements by compiling a sequence of basic tasks called laboratory unit operations (LUOs). Typical LUOs include weighing, grinding, manipulating, conditioning and separating. A particular analysis usually involves three to five operations, followed by a second analysis that may require different operations in a different order. Figure 1 outlines the organization and functions of LUOs.

To be truly effective, sample prepara-

Fig. 2 A laboratory robotic system, the Zymark DL 40. Robotic 'hands' are interchangeable to increase the instrument's flexibility.



tion should achieve this versatility without frequent manual intervention. To be economical as well, the instrument should adopt existing manual procedures and interface directly with existing analytical components.

Contemporary designs

A microprocessor-controlled robotic system is shown in Fig. 2. The robot arm can move up and down, in and out, and rotate 360 degrees within a cylinder of space centred around it. On the end of the arm is a 'hand' for picking up vials and test tubes. A real advantage in robotic sample preparation is the ability to change hands to accommodate containers, pipettes or temperature and pH probes. The steps in

a sample preparation procedure are performed at stations placed within reach of the robot arm. When the robotic arm has completed the steps (weighing, centrifuging, dispensing, heating), it either introduces the sample directly into the analytical instrument, or places it in a rack or carousel for subsequent analysis.

Microcomputers manage the interaction of robot, laboratory stations and the analytical instrumentation. The user controls the robot's motion via a keyboard entry. The microcomputer also acquires and processes data during the preparation, which enables it to revise or repeat procedures when necessary.

The robotics advantage

The robotic approach makes possible significant improvements in laboratory operations. For example, analytical precision is improved by eliminating human inconsistencies inevitably introduced in repetitive manual tasks. Each sample is prepared exactly according to the programmed procedure.

With automation, replicate samples, standards and controls can be prepared routinely and at low cost. Several sample preparation procedures may be programmed for around-the-clock operation, so that sample preparation and instrumental analysis can be more efficiently integrated. The physical set-up is convenient, compact and easy to change; often different methods can be performed within a single set-up, reducing the delays involved in multiple analysis. Methods development also benefits because of the ease of running multiple experiments to identify optimal conditions and error sources.

LUO class	Definition	Examples
Weighing	Quantitative measurement of sample mass	Direct measurement on balance
Grinding	Reducing sample particle size	Homogenizing
Manipulation	Physical handling of laboratory materials	Pouring Capping and uncapping containers Sample movement
Liquid handling	All physical handling of liquids - reagents & samples	Dispense, Dilute & Pipette Transfer (pumping & valving)
Conditioning	Modifying and controlling the sample environment	Time (start & stop) Temperature (heat & cool) Atmosphere (vacuum & gas blanket) Agitation (mix, stir, vortex & shake)
Measurement	Direct measurement of physical properties	pH, Conductivity etc. Absorbance, Fluorescence, etc.
Separation	Coarse mechanical and precision separations	Filtration - all techniques Partition - liquid-liquid & liquid-solid Centrifugation Precipitation Distillation & Recrystallization Electrophoresis
Control	Use of calculation and logical decisions in laboratory procedures	Calculate and dispense reagent to dilute to concentration based on weight or volume Adjust sample to desired pH

Fig. 1 Examples of laboratory unit operations (LUOs). Most analyses involve a combination of three to five LUOs.

Using robotics reduces human exposure to toxic or radioactive materials and harsh environments. Also, laboratory productivity is improved by freeing trained technicians from routine tasks. The electronic and mechanical modularity of a robotic system allows new laboratory stations to be added as user requirements change or as new modules are developed.

Automation prospects

A typical robotic sequence for chromatographic or spectroscopic analysis would follow these steps: weigh the sample and record the weight; add solvent to dissolve the sample; mix sample thoroughly; filter the sample solution and place an aliquot into the chromatograph or spectroscopy instrument. Robotics interfaces are available for gas chromatographs, liquid chromatographs and many types of spectroscopy instruments such as infrared, ultraviolet/visible, atomic absorption, and inductively coupled plasma spectroscopy. This procedure, outlined in Fig. 3, requires no human intervention. In addition, the robotic system can analyse the data and 'decide' whether it is necessary to reinject the sample or a standard, reconstruct a new calibration curve, or dilute the sample and reinject if it is outside the calibration range.

Other important applications include automated titrations and drug level determinations for metabolism studies. A robot-

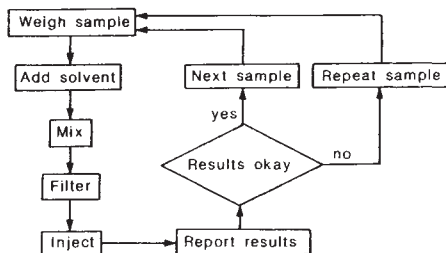


Fig. 3 A flow chart for common automated analysis.

ic system can also be used as an autosampler for techniques such as differential scanning calorimetry and superconducting nuclear magnetic resonance, and for physical testing instrumentation (measuring tensile strength, burst and so on). Automated microbiological procedures also offer fertile frontiers for laboratory robotics.

Robotic sample preparation is a young but rapidly expanding facet of the drive towards the totally automated laboratory. In less than three years, Zymark Corporation alone installed over 600 systems in the United States. Many potential uses for laboratory robotics systems have yet to be exploited, but it seems safe to assume that in the future these systems will make life in the laboratory a little easier. □

James N. Little is at Zymark Corporation, Zymark Center, Hopkinton, Massachusetts 01748, USA.

Portents of Pittsburgh 1986

High tech will be the talk of the town next week in Atlantic City. Here's a preview of the Pittcon display

Six years in the making, **supercritical fluid chromatography (SFC)** will arrive on the commercial scene via Lee Scientific and this spring's Pittsburgh Conference in Atlantic City (Reader Service No. 100). The new method achieves high resolution separation by pressure programming of fluids near and beyond their critical points. The supercritical fluids — gases that are compressed to near liquid densities — offer liquid-like solubility capabilities and gas-like flow properties, compatible with a wide range of detectors including ultraviolet absorbance, fluorescence and thermal conductivity systems. Lee lists high efficiency and fast analysis as advantages that its \$34,000 (US) product holds over more conventional liquid chromatography. Gas chromatographers may also appreciate the benefits of SFC in separating polar, non-volatile or thermally unstable compounds. The SFC 501 will be on display in booths 26011 and 27010.

The Ultrafinity-EP (epoxide) column combines optimal flow and chromatographic characteristics to make **affinity separations** up to 200 times faster than soft gel techniques (Reader Service No. 101). Beckman Instruments, Inc. will show this column and other fast affinity chromatography components at booths 723 to 737, and 822 to 836. Priced at \$365 (US) for the analytical units, Ultrafinity columns will isolate and purify antibodies, immunoglobulins, receptor proteins and enzymes.

Waters will present a new **single pump gradient HPLC** module at Pittsburgh, complete with multi-solvent delivery and programmable controller (Reader Service No. 102). The 600 gradient unit incorporates random phase synchronization soft-

ware for valve control that gives accurate (better than ± 0.5 per cent) and precise



Good grades on gradients for Waters's HPLC pump

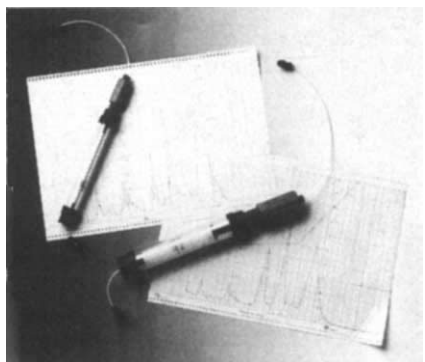
(better than ± 0.15 per cent) compositional proportioning. The Waters 600 can perform automatic start-up and shutdown without supervision, and takes up less than a foot of bench space. The module can be seen at booths 7001 to 7019 (odd), 8000 to 8014 (even), 8016 to 8019, 9016 to 9019, and 10014 to 10018 (even), and bought for \$13,100 (US).

The selectivity of affinity chromatography and the speed, resolution and sensitivity of HPLC come together in Pierce Chemical Co.'s **high performance liquid affinity chromatography (HPLAC)** columns (Reader Service No. 103). Available in boronate, concanavalin A and activated tresyl forms, SelectiSpher-10 columns cost between \$300 and \$450 (US) each. Pierce supplies the columns in 5, 10 and 25 cm lengths with 5 mm internal diameters. Stop by booths 7012 to 7018 (even) for more information.

Prepacked preparative high performance columns for **reversed phase chromatography** make analytical scale-ups easy.



Lee's basic SFC unit includes a flame ionization detector, capillary column and an IBM AT enhanced microcomputer



Pharmacia scales up chromatography for proteins and peptides

Pharmacia's PepRPC 15 μm and ProRPC 15 μm have selectivities comparable to their analytical counterparts, but their particle size is three times larger (Reader Service No. 104). Prices vary with column sizes (8 and 20 ml) and media; the PepRPC has 100 Å pore size silica coupled with C_2 and C_{18} chains, and the ProRPC is based on more porous silica (300 Å) to which C_2 and C_{18} chains are also attached. Pharmacia will be exhibiting at booths 4005 to 4009 at Pittcon.

Rheodyne markets **large-bore valves** designed to reduce the pressure drops encountered in analytical liquid chromatography by as much as 90 per cent (Reader Service No. 105). The valves accommodate flow rates over 100 ml per minute, with internal diameters measuring between 0.024 and 0.04 inches. Costs vary with valve variety; more detailed descriptions can be gathered at Rheodyne Inc.'s booths 655 through 659 (odd).

Expanding potential

The Teknivent Vector/One data system provides complete **data acquisition and digital control** for both mass spectrometers and gas chromatographs (Reader Service No. 106). Teknivent's system connects to any digital Hewlett-Packard GC/MS instrument for interactive tuning display, simultaneous data acquisition and processing and high speed graphics display. A \$14,500 (US) kit contains the interface electronics and PCSPEC executive software, but accompanying hardware is also available. The Vector/One package operates on any IBM PC, XT or AT personal computer under standard DOS 2.x or 3.x. Teknivent will be at Pittcon booths 238 and 240.

Houston Instruments plan to make **colour graphics** affordable with their \$699 (US) PC plotter (Reader Service No. 107). Built-in DM/PL intelligence makes the plotter compatible with most commercially available software packages. The plotter's four-pen carousel can compose vibrant charts, graphs and illustrations on standard $8\frac{1}{2} \times 11$ inch paper. HI's plotter

will appear at Pittcon, in booths G14 and G15.

The first demonstration of Cyborg Corp.'s Discovery software to the research community will take place at Cyborg's Pittsburgh booth 10051. A second generation **data acquisition and analysis program**, Discovery can handle area-under-the-curve calculations for chromatography, smoothing for vibration analysis or simple instrument control (Reader Service No. 108). Every operation is



spreadsheets

menu-selected, and Discovery can store sequences of data acquisition operation for repeated use. At \$1,190 (US), the package runs on IBM personal computers and requires a hard disk.

Spreadsheet computer graphics **software for liquid scintillation counting** may be as easy as Lotus 1-2-3 with Beckman's Data Capture software (Reader Service No. 109). Both the Data Capture package and model LS 2801 counter (Reader Service No. 110) will put in an appearance at Beckman booths 723 through 737 and 822 through 836. Beckman's software will transfer and store data from the scintillation counter, convert it to a form Lotus 1-2-3 'understands', create programs of specified experiment variables and access

data via customized programs in BASIC, FORTRAN or Pascal. The \$750 (US) package must be used with IBM PC, XT and AT personal computers.

Analysis round-up

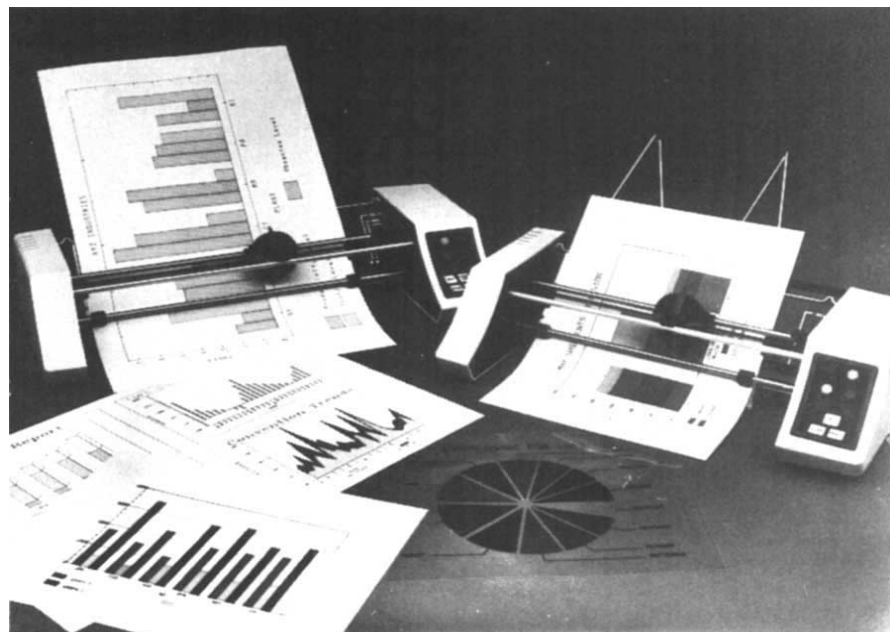
The 3370/3371 **X-ray fluorescent spectrometers** from Rigaku/USA Inc. are decked with DEC computers and normal or inverse optics (Reader Service No. 111). Rigaku's instrument features menu-driven software and full colour graphics for starters, with a price tag between \$175,000 and \$200,000 (US). Rigaku, at booths 944 to 951, have also developed their most complete **X-ray diffraction system** to date (Reader Service No. 112). A self-contained high-speed microcomputer provides for the simultaneous operation of two instruments, automatic detector parameter selection and computer-aided alignment. The D/Max-B series diffractometers sell for about \$70,000 (US), exclusive of data processing programs for



The 3370/3371 series spectrometer gets DEC system capabilities

search/match, quantitative and qualitative analysis, residual stress determination and other functions.

Dapple Systems provides **automation packages** for both X-ray diffractometers and X-ray spectrometers (Reader Service



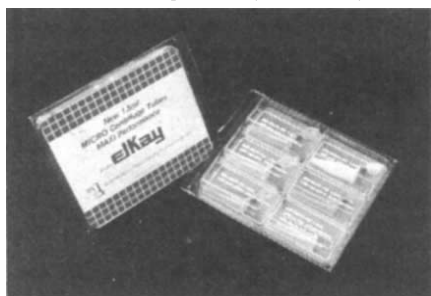
A PC plotter from Houston Instruments brightens any computer package

No. 113). Stepping motor assemblies allow precise control of the two-theta angle and a microcomputer takes care of analytical calculations. A complete \$25,000 (US) system consists of diffractometer or spectrometer, hardware, software, installation and training. Dapple will be doing business at Pittcon booth 19006.

For energy-intensive techniques such as photoacoustic spectroscopy, microscopy and diffuse reflectance, the FTS-40 at booth 8034 may steal the show (Reader Service No. 114). Bio-Rad's **spectrometer** boasts a research-grade optical bench, a powerful data processing system and a \$39,500 price tag. Bio-Rad will also have a protein microanalyser, a video densitometer and their monoclonal antibody purification system on display.

Spinning wheels

Elkay has **microcentrifuge tubes** in blue, green and yellow to avoid the confusion that can ruin a good day's work (Reader

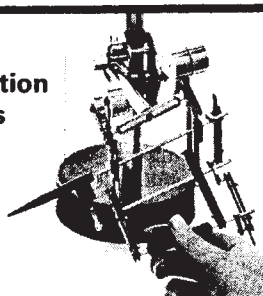


Elkay adds colour to centrifugation

Service No. 115). Their new 1.5 ml tubes are molded from polypropylene or polyethylene and fitted with a plug cap for a tight, spill-proof seal. Elkay's Pittcon booths 27031 and 27033 will have information on prices, which depend on size and composition.

ADVERTISEMENT

**Ideal for
microinjection
of xenopus
oocyte**



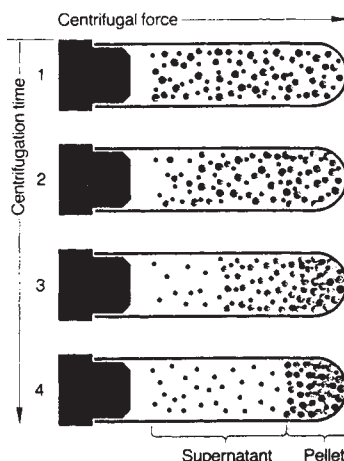
**SINGER MICROMANIPULATORS
ARE FAMOUS BECAUSE**

Singer Micromanipulators enable anyone to make exact movements without tremor or lost motion, in three dimensions, using a single control handle under low or high power magnification.

A varied range available. All simple to operate.

SINGER INSTRUMENT CO. LTD.,
Treborough Lodge, Roadwater, Somerset.
Tel: Washford (0984) 40226/325

Reader Service No.33



Beckman covers centrifugation from top to bottom

An array of **tabletop centrifuges and ultracentrifuges** will adorn Beckman's booths (Reader Service No. 116). The model L7 ultracentrifuge will make its debut in \$22,500 (US) water-cooled and \$23,450 (US) refrigerated versions, along with the AccuSpin tabletop and the variable speed Microfuge. Booths 723 through 737 and 822 through 836 will be loaded with product information for Beckman's line.

Tips on balances

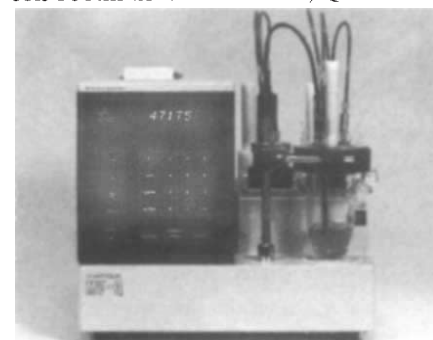
Analytical balances weigh in at Mettler's Pittcon booths. Their brochure, *Analytical and precision balances for the laboratory*, outlines over 20 models with charts and photographs for complete description (Reader Service No. 117). Mettler, at booths 8022 through 8029, can supply detailed answers to questions the brochure may provoke.

Brinkmann Instruments Co. offers a **top-loading analytical balance** with one-touch autocorrelation and an aerodynamic design to protect sample measurements from air currents (Reader Service No. 118). The \$3,095 (US) Sartorius model R160P features 0.1 second display update speed and a program menu that gives access to more than 38 functional settings. Operating parameters can change to suit the weighing environment, in digital filter settings, stability ranges, display format and zero tracking. Brinkmann is stationed at the Pittsburgh Conference in booths 6000 through 6015.

Innovative images

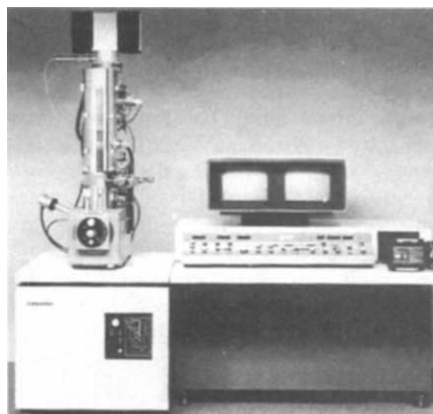
Quintel Corp.'s pack of Karl Fischer titrators is lead by the model MC-3 **coulometric titrator**, paired with the NE-1 moisture extractor in a \$12,000 (US) package (Reader Service No. 119). The MC-3 can analyse 10 µg of water in 20 seconds or 1 mg in 60 seconds, with a sample range of 0.1 µg to 10 mg and accuracy to 0.0001 per cent. Pyridine-free reagent sampling pro-

vides fast reactions, sharp end-point determinations and long shelf life. At Pittcon booths 23021 and 23023, Quintel can



The MC-3 titrator automatically monitors electrolyte voltage and background moisture level also demonstrate the instrument's internal microprocessor and software.

International Scientific Instruments Inc. is focusing on their **scanning electron microscope DS-130C**, which gives high resolution images even at low accelerating voltages (Reader Service No. 120). ISI has redesigned the objective lenses and



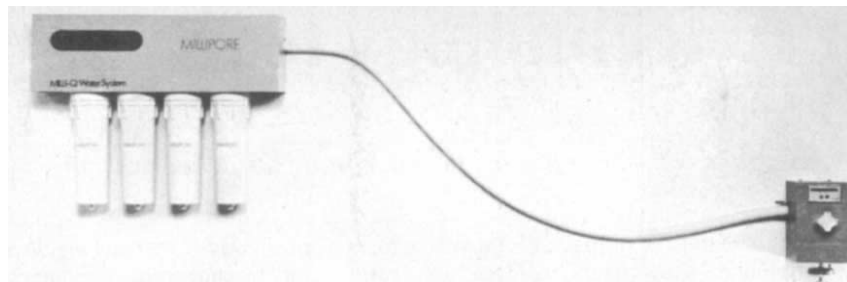
ISI's scanning electron microscope has an S-zone lens in its top stage

optical column so that the LaB₆ emitter provides 60 Å resolution at 1 kV, 120 Å at 500 kV and better than 20 Å resolution at 40 kV. The \$128,000 (US) machine's sample coverage is 6 inches, with 8 inches coverage optional. ISI will be exhibiting at booths 30019–30021 in Atlantic City.

Using **photon correlation spectroscopy**, Malvern Instruments' Autosizer IIc measures particles from 3 nm to 3 µm in



No particles escape unsized with Malvern's Autosizer IIc



Millipore provided a replacement cartridge kit and schedule to supplement its new water system

materials such as micelles, polymer latexes and microemulsions (Reader Service No. 121). Complete analysis on the **particle sizer** typically takes less than 30 seconds, and can accommodate all types of size distribution, whether narrow or broad, single- or multi-modal. The \$26,000 (US) Autosizer includes an 8-bit, 72-channel correlator and a 16-bit computer, with capabilities for diffusion coefficient and polymer molecular weight measurements. Malvern will show their complete range of sizing instruments at booths 249 and 251 at Pittcon.

Pittsburgh potpourri

The Vacuumate **vacuum pump** has a pull of 27 inches of mercury but runs as quiet as a whisper. Wheaton's pump can be hooked up with rotary vacuum evaporators and other laboratory equipment for high performance (Reader Service No. 122). Using compressed air and the Venturi



Wheaton's pump does its work in a vacuum

effect to create a vacuum, the non-electric pump has special benefits where explosion resistance is a pre-requisite, such as distilling solvents or other volatile liquids. For \$875 (US), the lightweight pump can also be sold with an optional foot pump for controlling input air pressure. Wheaton, in booths 204 to 212 (even), can furnish more information on the instrument.

Organic-free water or pyrogen-free water for bioresearch comes clean with Millipore Corp.'s new **water system**, the Milli-Q Plus (Reader Service No. 123). Automatic recirculation keeps water qual-

ity up to par between uses, and a colour-coded LED meter displays resistivity as water is being drawn. A Teflon outlet valve is fitted to protect against the risk of chemical contamination at the delivery point. The Milli-Q Plus is available in four-bowl and five-bowl units. At the Pittsburgh conference, Millipore will show the system at booths 9016 through 9018.

At booths 722 to 728 (even), Gilson Medical Electronics will display their new



The Medical Electronics FC 203 collects itself FC 203 **fraction collector** (Reader Service No. 124). The collector features multi-mode operation, a stationary rack system, and a footprint of less than one foot. Priced at \$1,190 (US), the FC 203 can adjust to a wide range of applications and offers the chromatographer a choice of six modes including time, drop and peak. Key parameters can also be changed with in-run editing.

Wherever a **low voltage/high current power supply** is needed, Haake Buchler Instruments Inc. has a 250 volt unit that fits the bill (Reader Service No. 125). Their Ephortec power supply provides 1.5 amps of current at voltages ideal for electrophoretic transfers of proteins or nucleic acids. An automatic switch prevents sample overheating and equipment damage by changing from regulated voltage to regulated current if maximum current is reached during a run. Haake Buchler, at Pittcon booths 1023 through 1031 (odd) and 2022 through 2030 (even), price the instrument at \$755 (US). □

These notes are compiled by Karen Wright from information provided by the manufacturers. To obtain further details about these products use the reader service card bound inside the journal.

ADVERTISEMENT

The two most notable high speed events of the year.

1. Halley's comet.
2. Protein separations in seconds.



Bio-Rad's new Microanalyzer™ cartridges perform ion exchange chromatography of protein samples with unprecedented speed. Their short format and an entirely new proprietary rigid polymeric packing allow flow rates as high as 5 ml/min and short elution times. This assures the high throughput needed for QC and biomedical applications.

These non-porous beads separate proteins of any molecular weight. Furthermore, the high recovery achieved with this packing makes these cartridges uniquely suited for micropreparative uses. Recoveries are normally greater than 95% with µg samples.

Three MA7 cartridges are exclusively available from Bio-Rad, including one for peptide analysis. Optimal performance is achieved by using these cartridges with our Protein Microanalyzer System. For full details, contact:

BIO-RAD

1414 Harbour Way South
Richmond, CA 94804
(415) 232-7000
800-4-BIO-RAD

Reader Service No.9

Answers.

Switching over to technology teaching

from Richard Pearson

The teaching of science and technological subjects in schools must be made more attractive and employers must do what they can to help.

THE graduate labour market is booming again. The *Employment* column last year reported that a record number of graduates went into jobs in 1984, demand grew again in 1985, and vacancies are expected to be a further 5–10 per cent higher in 1986. Graduate unemployment has fallen for the fourth year running, and is now at 11.2 per cent for scientists (and 6.4 per cent for engineers) from the universities, although it was still at a worryingly high 17.3 per cent for scientists graduating from the polytechnics (Table 1). All the signs are that demand for graduates will continue to increase, yet higher education is well into a phase of retrenchment.

In terms of the numbers graduating, the full effects of the government-imposed cuts in university expenditure in 1981 are only now beginning to filter through. At the time, the universities relied heavily on "natural wastage" and voluntary redun-

Table 1 Unemployment six months after graduation

	Universities	Polytechnics
Scientists	11.2%	17.3%
Engineers	6.4%	7.5%
All students	9.6%	13.6%

dancies. As a result, losses of academic staff were disproportionately higher than of support staff, particularly in subjects such as engineering and management where alternative employment was most readily available. This, together with the high equipment costs and uncertain student demand, hit the engineering courses particularly hard. Course numbers were cut and in 1985 the output of home students in the supposedly important technological subjects fell; in the case of electrical/electronic engineering output was down by 8.4 per cent, to be followed by a further 5.4 per cent fall this year. In computer science, numbers also fell in 1985 by 9 per cent. Admissions started picking up again in 1984 as selective funding was applied so the output in these, and in virtually every other subject, except physics and geology, will start to rise again in 1987, overall university numbers going up by about 7 per cent (see Fig. 1). There has been continuing growth in the polytechnics and other public sector institutions which has compensated in part for these falls, but as the polytechnics still account for only about 30 per cent of the total numbers graduating, the net effect is more one of static overall output¹. In the case of information technology (IT), how-

ever, the initiative embracing one-year postgraduate "conversion" courses for graduates in non-IT subjects has been a major boost.

The consultation period on the government's green paper in higher education is now over. If there is no major change in the government's stance, the next decade is likely to see a long-term decline in the output of higher education because of a continuing squeeze on finance. Engineering and technology are to be protected, however, the "switch" in this direction in 1985 resulting in a £43 million cash injection to increase the output in these subjects by about 1,500 a year from 1989.

While the government has been heavily criticized for the low level of its projections of student demand between now and 1995, it could be over-optimistic about the supply of suitable students to fill the planned 30 per cent increase in engineering and technology places unless some radical changes are quickly implemented.

There are two major concerns. The first is a possible shortage of students with mathematics and physics A levels. There will need to be a major growth in student interest in the information technology-related courses, especially from girls, who at present account for only about one in ten of the applicants, if the increased number of places are to be filled with high quality students. Between 1979 and 1984, the number of students passing A-level physics grew by 15 per cent and those passing A-level maths went up by 24 per cent. Yet the polytechnics are struggling to fill their science and technology places, with a number of empty places remaining in 1985, and the number of applications for university physics places fell in 1983 and again in 1984, possibly due to students switching to electronics and computer science courses. Between 1984 and the end of the decade, the number of students with two or more A-levels is expected to fall by about eight per cent due to demographic changes. It seems unlikely therefore that there will be the major increase needed in passes in mathematics and physics to fill the extra places. In addition, a repeat of the well-publicized problems of the electronics industry in 1985 will hardly encourage students to apply for electronics-related courses.

One of the causes of the already low standards of mathematics and physics achievements in schools is the shortage of suitable teachers. This will get worse over the next decade if science graduates con-

tinue to turn away from teaching. Applications for teacher-training courses in these subjects have fallen by over 20 per cent in the past year, and continuing low pay and teacher unrest are no incentive for the numbers to rise. Shortages of experienced lecturers in higher education are also likely to be a continuing problem as long as salaries and conditions remain uncompetitive.

Action is needed now to make teaching of these subjects more attractive. Dif-

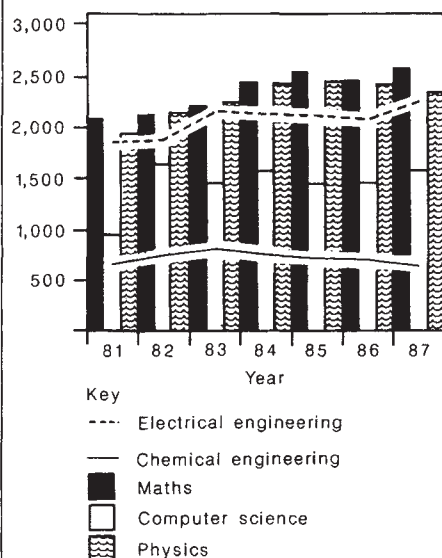


Fig. 1 Output of UK university graduates 1979–87, for home students only. Data for 1985–87 are projected. (Source: *Graduate Supply and Availability to 1987 and Beyond*, by R. Pearson (Institute of Manpower Studies, Brighton, 1986).

ferential salary increases for staff are likely to be far more effective than supplementary grants for students on teacher-training courses. If employers are to meet their future needs for manpower, they will have to become much more closely involved in collaborative arrangements such as sponsorship in order to attract a share of an inadequate supply. More importantly, employers will have to help encourage more able students to study science in schools and to enter vocationally relevant courses, and to support the teaching profession if shortages are not to become a permanent feature of the labour market over the next decade. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Brighton BN1 9RF, UK